Estas son traducciones al español de algunas de nuestras mejores obras que tienen gran éxito en todo el mundo. Editorial Prentice/Hall Internacional se complace en ponerlas al alcance del estudiante de habla española con lo que espera contribuir en gran medida a su mejor preparación profesional.

EDITORIAL PRENTICE/HALL INTERNACIONAL

0-13-112375-0

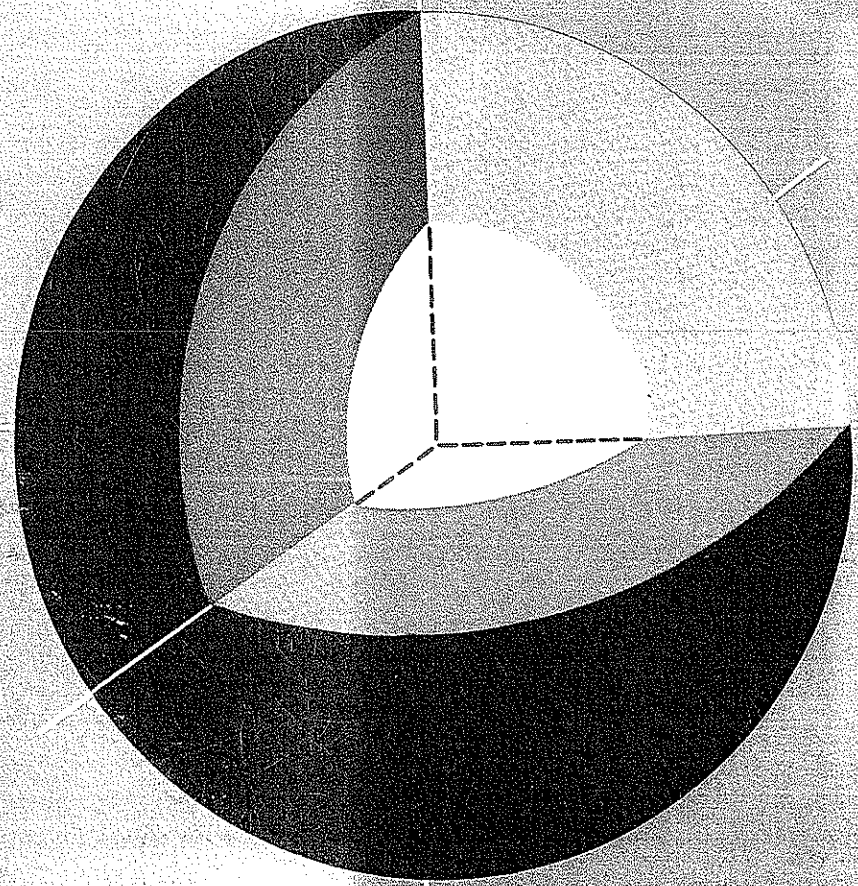
WILLIAMSON
CROWELL
TROTTER

Cálculo de funciones vectoriales

Selecciones

cálculo de funciones vectoriales

WILLIAMSON / CROWELL / TROTTER



PRENTICE/HALL INTERNACIONAL

Otros libros sobre MATEMATICAS PURAS de la PHI *en español* y de interés para usted:

ALGEBRA LINEAL

Por Kenneth Hoffman y Ray Kunze

Constituye el análisis más completo del álgebra lineal moderna que existe en la actualidad. Incluye más ecuaciones de primer grado, más cálculos con matrices, así como más problemas elementales y material elemental en general que otros textos que tratan sobre el mismo tema. Es un libro de texto fundamental para iniciarse en matemáticas a nivel universitario.

INTRODUCCION A LA TEORIA DE LOS NUMEROS

Por Anthony J. Pettofrezzo y Donald R. Byrkit

Describe tres áreas principales de estudio en la teoría de los números: 1) las propiedades de divisibilidad de los enteros, 2) la teoría de las congruencias, y 3) las fracciones continuas. Los temas se han elegido por su significación general para las matemáticas contemporáneas. Los autores incluyen también un breve estudio de las soluciones de las ecuaciones diofánticas.

CALCULO DE FUNCIONES VECTORIALES

La Editorial Prentice-Hall Internacional ha consagrado inusitada atención al aspecto lingüístico de los textos que, traducidos, ofrece al lector de lengua española. El lector puede estar seguro que tiene en sus manos una versión impecable de esta obra.

ASESORES TECNICOS

Carlo Federici Casa
Dr. en Física y Matemáticas
Universidad de Génova, Italia
Profesor de Matemáticas
Universidad Nacional de Colombia, Bogotá

Dr. César A. Trejo
Profesor de Matemáticas
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires, Argentina

Dr. Raúl Bravo Flores
Profesor de Matemáticas
Universidad de Chile, Santiago

ASESOR LINGÜISTICO
Dr. Carlos Vega
Profesor de Filología española

Cálculo de Funciones Vectoriales

RICHARD E. WILLIAMSON
Department of Mathematics
Dartmouth College

RICHARD H. CROWELL
Department of Mathematics
Dartmouth College

HALE F. TROTTER
Department of Mathematics
Princeton University

TRADUCCION Y ADAPTACION

Dr. NICOLAS YUS S.
Profesor de Matemáticas
Facultad de Ciencias
Universidad de Chile, Santiago
M. S. Columbia University U.S.A.
Ph. D. Brandeis University U.S.A.

PHI **EDITORIAL PRENTICE/HALL INTERNACIONAL**

Bogotá Buenos Aires Madrid México

CALCULO DE FUNCIONES VECTORIALES

Copyright © 1973 por Editorial Prentice-Hall Internacional, Englewood Cliffs, New Jersey. Reservados todos los derechos. Queda terminantemente prohibido reproducir este libro total o parcialmente sin permiso expreso de los editores.

Edición original en inglés, *Calculus of Vector Functions*, tercera edición, publicada en los Estados Unidos de América.
Copyright © 1972, Prentice-Hall, Inc. Englewood Cliffs. New Jersey 07632, U.S.A. Reservados todos los derechos. Es propiedad.

ISBN 0-13-112375-0

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*

Esta edición ha sido preparada por
Ediciones del Castillo, S. A.
Rodríguez San Pedro, 8 – Madrid-15
Depósito legal: M. 34.198-1972
Estampaciones Gráficas Grefol
Avda. Pedro Diez, 16 – Madrid-19

Impreso en España – Printed in Spain

Preámbulo

Ha sido el propósito de cada edición subsiguiente de este libro presentar el álgebra lineal y cálculo de tal modo que se complementen y realcen, pero manteniendo la esencia de cada tema como una disciplina separada. La presente edición representa nuestra mejor estimación de lo que actualmente consideramos un buen equilibrio de énfasis entre ambas materias como texto para muchos cursos en Universidades Estadounidenses. Esperamos que esta presentación sea adecuada para otras universidades con diferentes programas de estudio. La traducción al español por el profesor Yus es clara y precisa; y le quedamos muy agradecidos por su realización y a Prentice/Hall Internacional por su publicación.

R. E. WILLIAMSON
R. H. CROWELL
H. F. TROTTER

Prólogo

El presente libro es una introducción al cálculo de funciones de varias variables y al cálculo vectorial, bajo el punto de vista de que el cálculo consiste en la aproximación lineal de funciones. El único prerrequisito es el conocimiento del cálculo en una variable. El álgebra lineal necesaria se introduce en los dos primeros capítulos.

En la tercera edición se ha puesto énfasis en el aprendizaje y la comprensión de las técnicas matemáticas y de sus aplicaciones a problemas específicos. El álgebra lineal se usa para clarificar los conceptos y para enfatizar la utilidad de las técnicas algebraicas. Hemos dado todas las definiciones y teoremas en forma precisa, incluyendo, además, demostraciones completas de casi todas las afirmaciones dadas en el libro. Sin embargo, se supone que las demostraciones serán estudiadas sólo cuando ayudan a la comprensión. Si se estudian cuidadosamente los ejemplos, se puede sacar mucho provecho del libro sin estudiar las demostraciones. El libro ha sido diseñado de modo que sea lo más flexible posible permitiendo así que el profesor elija el nivel teórico más adecuado a los intereses y habilidades de los alumnos. Es poco probable que alguien elija uno de los dos extremos posibles: cubrir todas las demostraciones o no cubrir ninguna.

En la presente edición, la parte de álgebra lineal ha sido ampliada, dividiéndola en dos capítulos. El primero trata vectores y funciones lineales en $\mathcal{R}^n$ y contiene casi todas las ideas algebraicas básicas que se usan en el cálculo. El Cap. 2 contiene una discusión completa de la solución de sistemas lineales con una sección sobre aplicaciones. Presenta, además, las ideas fundamentales de espacio vectorial abstracto y dimensión y una sección breve pero completa

sobre las ecuaciones diferenciales con coeficientes constantes, para cuyo mejor estudio se incluye una sección sobre los espacios vectoriales complejos. También se tratan brevemente tópicos, tales como bases ortonormales y vectores propios.

En varias partes se ha incluido material relacionado con el cálculo numérico. Hay secciones sobre el método de Newton para funciones de varias variables, sobre estimación numérica de funciones definidas implícitamente y sobre la aproximación numérica de integrales definidas. Este material está adaptado al uso de computadores automáticos. En estas secciones hemos propuesto ejercicios que no pueden ser resueltos en un tiempo razonable sin uso de un computador, requiriéndose para ello programas bastante simples. Hemos incluido también ejercicios que no requieren el uso de un computador.

Queremos agradecer a varias personas en varias universidades por sus sugerencias sobre cómo mejorar el libro. Su interés y ayuda ha sido muy bienvenida. También agradecemos a la señora Dorothy Krieger y al señor Arthur Wester, de Prentice-Hall, cuyos esfuerzos han ayudado a hacer más atractivo y útil este libro.

R. E. WILLIAMSON
R. H. CROWELL
H. F. TROTTER

Prólogo del traductor

La traducción está lista y el editor me pide un prefacio. Pasa el tiempo y no sé qué decir. Los méritos del libro son claros, como se podrá comprobar fácilmente echándole una ojeada. En todo caso, el lector probablemente no tome muy en cuenta los comentarios que al respecto haga el traductor (si es que lee este prefacio). Por tanto, me parece más apropiado hacer algunos comentarios acerca de la traducción misma.

Para todos es clara la necesidad de que nuestros estudiantes dispongan de textos en español y estarán de acuerdo en agradecer los esfuerzos de Prentice-Hall para proporcionarlos.

Espero que la traducción que he hecho no adolezca de los defectos que he criticado duramente en otros. No creo haber incurrido en los errores de conceptos comunes en textos traducidos por personas que desconocen el tema. Sin embargo, estoy seguro de que, a pesar de todos mis esfuerzos, he usado términos y expresiones que, a pesar de ser las habituales en mi país, puedan ser desagradables o aún inentendibles para los lectores de otros lugares. Un esfuerzo para soportar y perdonar esta situación es el precio que debemos pagar por el privilegio de tener obras matemáticas en castellano (o español, como se dice en otros países). La única alternativa es desarrollar un lenguaje matemático común, pero para ello se necesitarían muchos más contactos que los que son habituales ahora o que se pueden prever para el futuro.

Santiago de Chile, julio de 1972

NICOLÁS YUS S.

Programas posibles

Hemos tratado de organizar el libro de modo que cada sección conduzca naturalmente a la siguiente, sin embargo, no es necesario incluir todas las secciones ni tomarlas en el orden dado. En particular, no es necesario completar el álgebra lineal antes de empezar el cálculo (los estudiantes que hayan visto el álgebra lineal podrán empezar por el Cap. 3 y usar los dos primeros capítulos como repaso y referencia). Lo esencial para los Caps. 3 a 7 está contenido en las siete primeras secciones del Cap. 1. (Una excepción es el uso de algunos resultados acerca de la dimensión en la última sección del Cap. 4.) El estudio de los determinantes puede ser postergado hasta la sección sobre el cambio de variables en integrales múltiples y las últimas tres secciones sobre análisis vectorial. Por otra parte, en la práctica pueden usarse los determinantes para invertir matrices pequeñas (2-por-2 o 3-por-3) o para resolver sistemas de pocas ecuaciones lineales, de modo que algunos podrán estudiar determinantes primero y posponer (u omitir) el método de reducción por filas dado en la primera sección del Capítulo 2. También se puede restringir la discusión de la inversión de matrices al caso trivial 2-por-2 a costa de omitir la consideración más tarde, de algunos ejercicios en dimensiones mayores.

Tampoco hay dificultades en hacer otros cambios en el orden, como, por ejemplo, estudiar más pronto las integrales múltiples.

Las sucesiones de números de secciones que siguen a continuación son mínimas en el sentido de que ellas contienen los prerrequisitos mínimos para el material considerado más adelante sin incluir todo lo que sería conveniente cubrir en un curso típico. Más secciones pueden ser agregadas de acuerdo con las necesidades para un curso de un año. La experiencia ha mostrado que en

un curso se deben usar dos periodos por sección como término medio. Algunas más cortas solo requieren una y otras más largas, tres.

Cap. 1	*Cap. 2*	*Cap. 3*	*Cap. 4*	*Cap. 5*	*Cap. 6*	*Cap. 7*	*Comentarios*
1–5, 7		1–8	1–3, 5	1, 5	1–3	1, 3–5	Vista panorámica
1–5, 7, 8	1	1–9	1–3, 5	1	1–3, 6		Orientación numérica
1–5, 7	4, 5	1–8	1–3, 5, 7		1–4		Enfasis en la dimensión
1–5, 7	4, 5, 8, 9	1–8	1–3	5–7			Sin integrales múltiples
1–5,7		1–8	1–3, 5	1	1–3		Curso semestral
1–5	1	1–8	1–3	1	1, 2, 7	1	Sin determinantes

Tabla de materias

Preámbulo v
Prólogo vii
Prólogo del traductor ix
Programas posibles xi

Introducción *1*

Capítulo 1. Vectores y linealidad *4*

1. Vectores 4
2. Interpretaciones geométricas 7
3. Matrices 15
4. Funciones lineales 24
5. Producto punto 35
6. Geometría euclideana 44
7. Determinantes 49
8. Desarrollo de determinantes 64

Capítulo 2. Algebra lineal *75*

1. Ecuaciones lineales, matrices inversas 75
2. Algunas aplicaciones 84
3. Teoría de ecuaciones lineales 92
4. Espacios vectoriales, subespacios, dimensiones 102
5. Funciones lineales 114

6. Operadores diferenciales 123
7. Espacios vectoriales complejos 139
8. Bases ortonormales 152
9. Vectores propios 159
10. Coordenadas 170

Capítulo 3. Funciones y derivadas *180*

1. Funciones vectoriales 180
2. Funciones de una variable 194
3. Longitud de arco 202
4. Integrales de línea 210
5. Derivadas parciales 217
6. Derivadas parciales vectoriales 223
7. Límites y continuidad 229
8. Diferenciales 240
9. El método de Newton 252

Capítulo 4. Cálculo vectorial *261*

1. Derivadas direccionales 261
2. El gradiente 266
3. La regla de la cadena 276
4. Funciones definidas implícitamente 290
5. Coordenadas curvilíneas 298
6. Teoremas de la función inversa y de la función implícita 313
7. Superficies y tangentes 325

Capítulo 5. Funciones reales *338*

1. Valores extremos 338
2. Polinomios cuadráticos 348
3. Desarrollos de Taylor 366
4. Desarrollos de Taylor y valores extremos 376
5. Series de Fourier 382
6. Desarrollo de Fourier modificado 390
7. Las ecuaciones del calor y de las ondas 396
8. Convergencia uniforme 406
9. Funciones ortogonales 413

Capítulo 6. Integrales múltiples *423*

1. Integrales iteradas 423
2. Integrales múltiples 432
3. Cambio de variables 449

4. Integrales impropias 463
5. Estimación de integrales 473
6. Integración numérica 479

Capítulo 7. Teoría de los campos vectoriales *490*

1. El teorema de Green 490
2. Campos vectoriales conservativos 503
3. Integrales de superficie 510
4. El teorema de Stokes 521
5. El teorema de Gauss 530
6. Los operadores ∇, ∇x y $\nabla\cdot$ 536
7. Formas diferenciales 544
8. La derivada exterior 553

Apéndice *557*

1. Introducción 557
2. La regla de la cadena 561
3. Fórmula del largo de arco 563
4. Convergencia de las series de Fourier 565
5. Demostración de los teoremas de la función inversa y de la función implícita 567
6. Demostración del teorema de Lagrange 573
7. Demostración del teorema de Taylor 575
8. Existencia de la interal de Riemann 578
9. La fórmula del cambio de variables para integrales 582

Indice *587*

Introducción

Un primer curso de cálculo trata sobre funciones de una variable con valores reales, es decir, funciones definidas en la recta real $\mathbb{R}$, o en una parte de ella, y con valores en $\mathbb{R}$. Por ejemplo, una fórmula como $y = x^2 + 3$ proporciona un número real y por cada número real x y define así una función f de $\mathbb{R}$ en $\mathbb{R}$ tal que (por ejemplo) $f(0) = 3$, $f(-2) = 7$, $f(\sqrt{3}) = 6$, etc. En este libro, nos preocuparemos de las funciones de varias variables y cuyos valores pueden ser números reales o, más generalmente, m-plas de números reales. Por ejemplo, un par de fórmulas:

$$y_1 = \sqrt{x_1^2 + x_2^2 + x_3^2}$$
$$y_2 = x_1 x_2 + 5x_3$$

proporciona un par de números (y_1, y_2) por cada triple (x_1, x_2, x_3) de números y define una función g del «espacio 3-dimensional» en el «espacio 2-dimensional». En particular, las fórmulas anteriores nos dan

$$g(0, 0, 0) = (0, 0),$$
$$g(1, 2, 3) = (\sqrt{14}, 17),$$
$$g(3, 2, 1) = (\sqrt{14}, 11).$$

Usaremos el símbolo $\mathbb{R}^n$ para denotrar el conjunto de todas las n-plas de números reales. ($\mathbb{R}^1$ es lo mismo que $\mathbb{R}$.) El **dominio** de una función es el con-

junto en el cual ella está definida y su **rango** o **imagen** es el conjunto de los valores asumidos por la función.

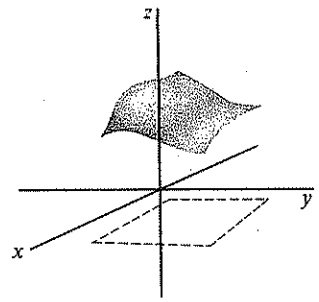
Hablamos de funciones «de $\mathcal{R}^n$ en $\mathcal{R}^m$» y escribimos

$$\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$$

para indicar que f es una función cuyo dominio es un subconjunto de $\mathcal{R}^n$ y cuya imagen es un conjunto de $\mathcal{R}^m$. $\mathcal{R}^n$ se llama el **espacio dominio** y $\mathcal{R}^m$ el **espacio imagen.** A veces se usan los términos «transformación» o «mapeo» en lugar de «función».

A pesar de que las funciones de una variable son básicas y muy útiles, hay muchas situaciones cuya formulación matemática requiere el uso de las funciones que estamos considerando. Por ejemplo, así como ciertas curvas en el plano pueden ser representadas como gráficos de funciones de $\mathcal{R}^1$ en $\mathcal{R}^1$, del mismo modo, algunas superficies en 3-espacio pueden ser representadas como gráficos de funciones de $\mathcal{R}^2$ en $\mathcal{R}^1$. La Fig. 1 muestra el gráfico de la función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^1$ definida por la fórmula:

$$f(x, y) = z.$$



Otros ejemplos de curvas y superficies que pueden ser descritas como funciones serán dadas al comienzo del Cap. 3.

La mayoría de los problemas y ejemplos de este libro tratarán sobre funciones de $\mathcal{R}^n$ en $\mathcal{R}^m$ con valores de m y n no mayores de 2 o 3, ya que los problemas en dimensiones mayores son difíciles de visualizar y a menudo requieren computaciones demasiado complicadas. Por el contrario, en el tratamiento teórico, formularemos los resultados para el caso de dimensiones arbitrarias. Esto no constituye una generalización vacía. Un economista podrá desear considerar un modelo matemático en el cual los precios de ciertos artículos son determinados por un número de variables, tales como costos de producción y demanda. Un ingeniero civil podrá querer estudiar los desplazamientos en las junturas de una estructura compleja para varias combinaciones de cargas aplicadas en varios puntos. Ahora que las computadoras automáticas han hecho posibles los cálculos aritméticos, se pueden atacar y resolver problemas

con docenas o aun varios cientos de variables. Por tanto, es importante disponer de técnicas y teoremas aplicables a funciones de $\mathcal{R}^n$ en $\mathcal{R}^m$ para todo valor de m y n.

La mayor parte de las ideas estudiadas en el cálculo de una variable reaparecen de una forma más general en el cálculo de varias variables. Por ejemplo, el problema de encontrar la recta tangente al gráfico de $y = f(x)$ de una función de una variable, se transforma en el problema de encontrar el plano tangente al gráfico de $z = g(x, y)$ de una función de dos variables. La recta tangente tiene una ecuación de la forma $y = mx + b$ y el coeficiente m se encuentra calculando la derivada de f. En el caso de dimensiones mayores el plano tangente tiene una ecuación de la forma $z = px + q + y + c$ y, como veremos, los coeficientes p y q están dados por la derivada de g en dimensiones mayores.

El cálculo diferencial puede ser descrito como la técnica para estudiar las funciones generales mediante la aproximación por funciones de formas más simples, tales como $y = mx + b$ o $z = px + qy + c$. En una dimensión, estas funciones «lineales» son muy simples, pero en dimensiones mayores, ellas son más complicadas. Los dos primeros capítulo del libro están dedicados a las funciones lineales y a otros temas cercanos que tienen aplicaciones en muchos otros campos fuera del cálculo.

El fin principal del libro es el estudio de las generalizaciones de la derivada, la integral a dimensiones más elevadas y los capítulos posteriores están dedicados a ello.

Estudiaremos las relaciones entre estas ideas y en cómo usarlas en la solución de una amplia gama de problemas. Las técnicas de derivación e integración del cálculo de una variable son aplicables al cálculo de varias variables como método de computación. Sin embargo, las interpretaciones y, a menudo las ideas básicas, son bastante distintas. Para desarrollar la intuición geométrica conviene aprender a visualizar los dibujos en tres dimensiones. Los de dos dimensiones son fáciles de dibujar y son más sencillos. Una gran ayuda para la comprensión de los problemas tridimensionales y su relación con el mundo físico es la habilidad para hacer dibujos en perspectiva.

Las letras en negrita $\mathbf{x}$, $\mathbf{a}$, etc., que se usan para distinguir los vectores de los números pueden ser escritas a mano de varias maneras. Algunas posibilidades son $\vec{x}$, $\tilde{x}$, $\underline{x}$. también se pueden usar mayúsculas siempre que no se las confunda con las notaciones habituales para matrices y números. La aparición en negrita de una palabra significa que ella está siendo definida en ese lugar.

1

Vectores y linealidad

1. Vectores

Denotaremos por $\mathcal{R}^n$ al conjunto de los n-plas $(x_1, x_2, \ldots, x_n)$ de números reales. Adoptaremos la convención de usar letras en negrita $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, etc., para denotar n-plas en $\mathcal{R}^n$ y letras ordinarias para designar números reales. En particular, podremos escribir $\mathbf{x} = (x, y)$ o $\mathbf{x} = (x, y, z)$ para denotar 2-plas o 3-plas genéricas para evitar el uso de subíndices.

Para dos elementos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ en $\mathcal{R}^n$, definimos su **suma** $\mathbf{x} + \mathbf{y}$ como el n-pla

$$(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Para cualquier número real r y n-pla $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, definimos el **múltiplo numérico** $r\mathbf{x}$ como la n-pla $(rx_1, rx_2, \ldots, rx_n)$. Por ejemplo, si $\mathbf{x} = (2, -1, 0, 3)$ e $\mathbf{y} = (0, 7, -2, 3)$, entonces

$$\mathbf{x} + \mathbf{y} = (2, 6, -2, 6)$$

y

$$3\mathbf{x} = (6, -3, 0, 9).$$

Escribimos $-\mathbf{x}$ para el múltiplo numérico $(-1)\mathbf{x}$ y $\mathbf{x} - \mathbf{y}$ como abreviación de $\mathbf{x} + (-\mathbf{y})$. Usamos 0 para denotar la n-pla que consiste enteramente de ceros. (Esta notación es ambigua, ya que, por ejemplo, 0 puede significar (0, 0) en una fórmula y (0, 0, 0) en otra. Esta ambigüedad raramente causa confusión, ya que en la mayor parte de los contextos solo una interpretación tiene sentido. Por ejemplo, si $\mathbf{z} = (-2, 0, 3)$, entonces en la expresión $\mathbf{z} + 0$, el 0 deberá

representar (0, 0, 0), ya que la suma está definida solo para n-plas con el mismo número de coordenadas.)

Las fórmulas que damos a continuación son válidas para arbitrarios $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$ en $\mathcal{R}^n$ y números r, s arbitrarios. Ellas expresan leyes para nuestras nuevas operaciones de suma y multiplicación numérica, que son muy parecidas a las conocidas leyes de distributividad, conmutatividad y asociatividad en el caso de la suma y multiplicación habitual de números.

1. $r\mathbf{x} + s\mathbf{x} = (r + s)\mathbf{x}$.
2. $r\mathbf{x} + r\mathbf{y} = r(\mathbf{x} + \mathbf{y})$.
3. $r(s\mathbf{x}) = (rs)\mathbf{x}$.
4. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
5. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
6. $\mathbf{x} + 0 = \mathbf{x}$.
7. $\mathbf{x} + (-\mathbf{x}) = 0$.

Estas leyes son consecuencias bastante obvias de las definiciones de estas operaciones y de las leyes de la aritmética. Como ilustración, daremos una demostración formal de la ley 2.

Sean $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ y sea r un número real. Entonces

$$r\mathbf{x} = (rx_1, rx_2, \ldots, rx_n) \qquad \text{[definición de multiplicación numérica]}$$

$$r\mathbf{y} = (ry_1, ry_2, \ldots, ry_n) \qquad \text{[definición de multiplicación numérica]}$$

y, por tanto,

$$r\mathbf{x} + r\mathbf{y} = (rx_1 + ry_1, rx_2 + ry_2, \ldots, rx_n + ry_n) \qquad \text{[definición de suma]}$$

Por otra parte,

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \qquad \text{[definición de suma]}$$

y, por tanto,

$$r(\mathbf{x} + \mathbf{y}) = (r(x_1 + y_1), r(x_2 + y_2), \ldots, r(x_n + y_n)) \qquad \text{[definición de multiplicación numérica]}$$

Por la ley de distributividad de la aritmética ordinaria, $r(x_1 + y_1) = rx_1 + ry_1$, $r(x_2 + y_2) = rx_2 + ry_2$, etc., y, por tanto, las n-plas $r\mathbf{x} + r\mathbf{y}$ y $r(\mathbf{x} + \mathbf{y})$ son iguales, como se quería probar.

Cualquier conjunto con operaciones de suma y multiplicación por números reales definidos de tal manera que se cumplan las leyes 1-7, se llama un **espacio vectorial** y sus elementos se llaman **vectores.** En el Cap. 2 veremos otros espacios

vectoriales; pero por ahora, se puede considerar que «vector» significa «elemento de $\mathcal{R}^n$» para algún n. Los números se llaman a veces **escalares** cuando se desea enfatizar la diferencia entre números y vectores. En física, por ejemplo, uno puede considerar la masa y la energía como cantidades escalares en contraposición a cantidades vectoriales, tales como velocidad o momento. El término múltiplo escalar es sinónimo de lo que hemos llamado múltiplo numérico.

Los vectores $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, \ldots, 0, 1)$ tienen la propiedad de que si $\mathbf{x} = (x_1, \ldots, x_n)$, entonces

$$\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n.$$

Debido a que todo elemento de $\mathcal{R}^n$ puede ser representado en forma única de esta manera, el conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de vectores se llama la base **natural** de $\mathcal{R}^n$. Por ejemplo, en $\mathcal{R}^3$ tenemos $(1, 2, -7) = \mathbf{e}_1 + 2\mathbf{e}_2 - 7\mathbf{e}_3$. Una suma $x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$ de múltiplos numéricos se llama una combinación lineal de los vectores $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Más generalmente, una suma $a_1\mathbf{x}_1 + \ldots + a_n\mathbf{x}_n$ de múltiplos se llama una **combinación lineal** de los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Así, por ejemplo, la ecuación

$$(2, 3, 4) = 4(1, 1, 1) - 1(1, 1, 0) - 1(1, 0, 0)$$

muestra que el vector (2, 3, 4) puede ser representado como una combinación lineal de los vectores (1, 1, 1), (1, 1, 0) y (1, 0, 0).

EJERCICIOS

1. Dados $\mathbf{x} = (3, -1, 0)$, $\mathbf{y} = (0, 1, 5)$ y $\mathbf{z} = (2, 5, -1)$ calcular $3\mathbf{x}$, $\mathbf{y} + \mathbf{z}$ y $4\mathbf{x} - 2\mathbf{y} + 3\mathbf{z}$. [*Resp.* (19, 9, −13).]

2. Encontrar números a y b tales que $a\mathbf{x} + b\mathbf{y} = (9, -1, 10)$, donde $\mathbf{x}$ e $\mathbf{y}$ son como en el Ejercicio 1. ¿Hay más de una solución?

3. Demostrar que ninguna selección de números a y b es tal que $a\mathbf{x} + b\mathbf{y} = (3, 0, 0)$, donde $\mathbf{x}$ e $\mathbf{y}$ son los del Ejercicio 1. ¿Para qué valor(es) de c (si existen) se satisface la ecuación $a\mathbf{x} + b\mathbf{y} = (3, 0, c)$?

4. Describa una demostración de las leyes 3 y 4 de la Sec. 1, dando una justificación precisa de cada paso.

5. Verifique que el conjunto $C[a, b]$ de todas las funciones continuas con valores definidos en el intervalo $a \leq x \leq b$ es un espacio vectorial, con la suma y multiplicación numérica definida por $(f + g)(x) = f(x) + g(x)$ y $(rf)(x) = rf(x)$.

6. Demuestre que la representación de un vector x de $\mathcal{R}^n$ con respecto a la base natural es única. Es decir, demuestre que si

$$x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = y_1\mathbf{e}_1 + \ldots + y_n\mathbf{e}_n,$$

entonces $x_k = y_k$ para $k = 1, \ldots, n$.

7. En los casos (a) y (b) dados más abajo, represente el primer vector como combinación lineal de los vectores restantes, ya sea por simple inspección o resolviendo un sistema adecuado de ecuaciones.
   (a) (2, 3, 4); (1, 1, 1), (1, 2, 1), (−1, 1, 2).
   (b) (2, −7); (1, 1), (1, −1).
   (c) (−2, 3); $\mathbf{e}_1$, $\mathbf{e}_2$.

2. Interpretaciones geométricas

Usando coordenadas, es posible representar $\mathcal{R}^1$ como una recta, $\mathcal{R}^2$ como un plano y $\mathcal{R}^3$ como un espacio 3-dimensional.

Para representar $\mathcal{R}^1$ como una recta, es preciso especificar un punto de ella, que será llamado el **origen**, una unidad de distancia, y un sentido en la recta que será llamado positivo (el sentido opuesto se llama entonces negativo). Entonces a un número positivo x le corresponde el punto que está a una distancia x, en el sentido positivo, del origen. A un número negativo x le corresponde el punto a una distancia $|x|$ del origen en el sentido negativo. Al número cero le corresponde el origen. La recta numérica se considera generalmente en forma horizontal con el sentido positivo hacia la derecha. Con esta convención estándar, obtenemos la acostumbrada Fig. 1, en la cual la flecha indica el sentido positivo.

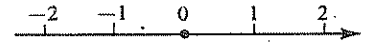


Figura 1

En el plano, se toman un origen, una unidad de distancia, dos rectas perpendiculares (llamadas **ejes**) que pasan por el origen y un sentido positivo en cada eje. Dado un vector de $\mathcal{R}^2$, es decir, un par (x_1, x_2) de números, el procedimiento del párrafo anterior hace corresponder al número x_1 un punto p_1 en el primer eje y el número x_2 un punto p_2 en el segundo eje. Entonces al vector (x_1, x_2) le corresponde el punto p del plano tal que su proyección en el primer eje es p_1 y su proyección en el segundo eje es p_2. La **proyección** (perpendicular) de un punto p en una recta L se define como el pie de la perpendicular desde p a L, si p no está en L. Si p está en L, entonces la proyección de p en L es simplemente p.

La convención habitual es de considerar el primer eje como horizontal con el sentido positivo hacia la derecha y el segundo eje como vertical con el sentido positivo hacia arriba. Esto conduce a la situación acostumbrada de la Fig. 2.

Si representamos un vector como una flecha desde el origen al punto correspondiente, como lo hemos hecho en la Fig. 2, se obtiene una figura mejor que si simplemente marcamos el punto.

Una extensión obvia de este procedimiento sirve en el caso de tres dimensiones. Se toman un origen, tres ejes perpendiculares, que pasan por él, y un

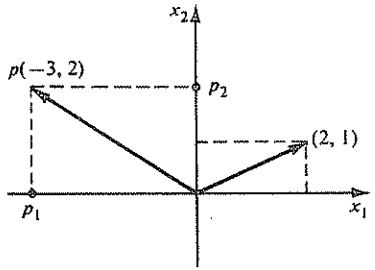


Figura 2

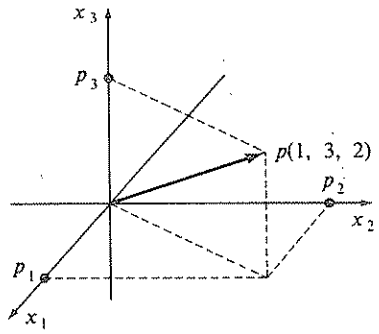


Figura 3

sentido positivo en cada uno de ellos. Un vector de $\mathcal{R}^3$ es una 3-pla (x_1, x_2, x_3) de números y a él se pueden asociar puntos p_1, p_2, p_3 en los ejes. Entonces al vector (x_1, x_2, x_3) le corresponde el punto p cuyas proyecciones en los ejes son p_1, p_2 y p_3.

No hay una convención universalmente aceptada para designar los ejes. La Fig. 3 ilustra la convención que se usará en este libro, sin embargo, hay otras convenciones de uso común.

Hemos descrito cómo hacer una correspondencia entre rectas (n-plas de números) y puntos. Ahora consideraremos la interpretación geométrica de las operaciones vectoriales de suma y multiplicación numérica. Ahora, representamos un vector como una flecha que va desde el origen al punto con las coordenadas dadas. La Fig. 4 muestra dos vectores $\mathbf{u} = (u_1, u_2)$ y $\mathbf{x} = (v_1, v_2)$. Por la definición de la suma en $\mathcal{R}^2$, $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2)$, de modo que la flecha que corresponde a $\mathbf{u} + \mathbf{v}$ debe ser dibujada como se indica allí.

Como lo sugiere la Fig. 4, $\mathbf{u} + \mathbf{v}$ está representado por una flecha que va del origen al vértice opuesto del paralelogramo cuyos lados son las flechas que representan a $\mathbf{u}$ y $\mathbf{v}$. Esta regla geométrica para sumar vectores se llama a menudo la **ley del paralelogramo para la suma.** (Por supuesto, para demostrar que esta ley es una consecuencia de nuestras definiciones, es necesario usar algunos teoremas de geometría. En el Ejercicio 4(a) se da un esquema de la demostración.)

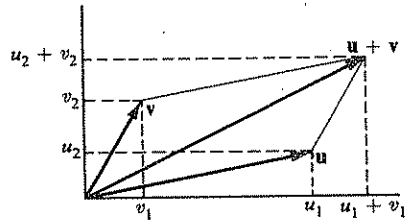


Figura 4

La Fig. 5 ilustra otro método para sumar vectores geométricamente. Partiendo del extremo de la flecha que representa a $\mathbf{u}$, se dibuja una flecha de igual largo y paralela a la flecha que representa a $\mathbf{v}$. (En otras palabras, se **traslada** la flecha que representa a $\mathbf{v}$ desde el origen al extremo de $\mathbf{u}$.) Entonces $\mathbf{u} + \mathbf{v}$ está representado por una flecha desde el origen hasta el extremo de la trasladada de $\mathbf{v}$. Esta regla puede ser aplicada a cualquier par de vectores mientras que, estrictamente hablando, la regla del paralelogramo no puede ser aplicada si las flechas que representan a $\mathbf{u}$ y $\mathbf{v}$ están en la misma recta.

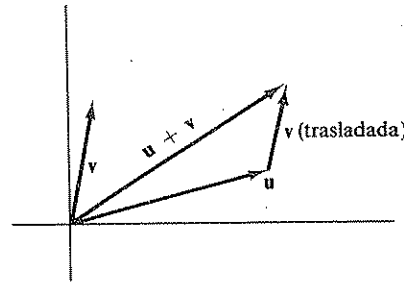


Figura 5

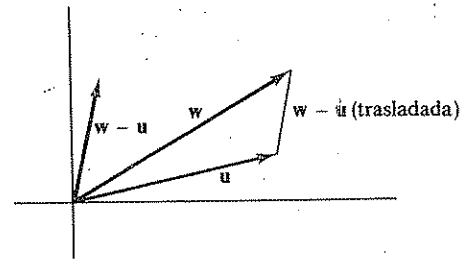
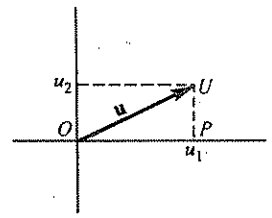


Figura 6

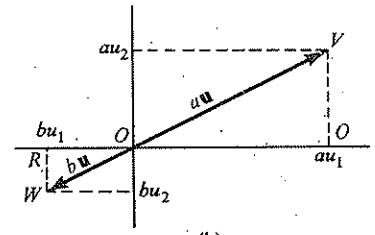
Si denotamos a $\mathbf{u} + \mathbf{v}$ por $\mathbf{w}$, entonces $\mathbf{v} = \mathbf{w} - \mathbf{u}$. La Fig. 6 (que es igual a la Fig. 5) ilustra el hecho útil de que la diferencia de dos vectores $\mathbf{w}$ y $\mathbf{u}$ está representada por una flecha desde el extremo de $\mathbf{u}$ al extremo de $\mathbf{w}$, trasladada apropiadamente.

La Fig. 7 ilustra la multiplicación numérica tanto por un número positivo a como por un número negativo b. Para un número positivo a, la flecha que representa $a\mathbf{u}$ tiene la misma dirección que la flecha que representa a $\mathbf{u}$ y es a veces más larga. Para un número negativo b, $b\mathbf{u}$ tiene la dirección opuesta de $\mathbf{u}$ y es $|b|$ veces más larga (véase Ejercicio 6(b)).

La decisión de representar geométricamente a un vector como un punto, $\cdot$, como una flecha o quizá como una flecha trasladada, depende de la situación geométrica en que aparezca el vector. Desde el punto de vista lógico, nos po-



(a)



(b)

Figura 7

dríamos restringir a usar solo puntos o solo flechas, evitando así el tener que tomar una decisión. Sin embargo, es muy deseable para la comprensión geométrica el poder usar ambas interpretaciones y los ejemplos mostrarán cuál es más útil en un problema particular.

Hasta ahora hemos usado informalmente rectas en $\mathcal{R}^2$ y $\mathcal{R}^3$ para obtener representaciones geométricas de la suma y la multiplicación numérica. Una vez hecho esto, podemos definir formalmente lo que entenderemos por una recta en $\mathcal{R}^n$. Para empezar, fijamos un vector $\mathbf{x}_1$ *distinto de cero* y llamamos recta por el origen al conjunto $\mathcal{L}_0$ de todos los múltiplos numéricos $t\mathbf{x}_1$. Entonces, definimos la **recta** $\mathcal{L}$ paralela a $\mathcal{L}_0$ como el conjunto de todos los puntos que pueden ser representados en la forma $t\mathbf{x}_1 + \mathbf{x}_0$, donde $\mathbf{x}_0$ es un vector fijo. La relación entre $\mathcal{L}_0$ y $\mathcal{L}$ se muestra en la Fig. 8, en el caso en que los vectores $\mathbf{x}_1$ y $\mathbf{x}_0$ están en $\mathcal{R}^3$.

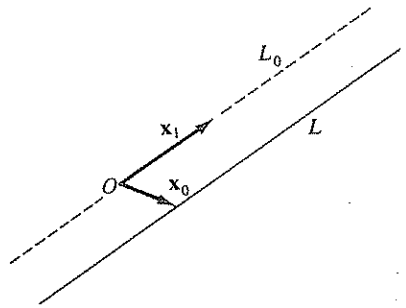


Figura 8

Decimos que dos rectas, una formada por todos los puntos de la forma $s\mathbf{u}_1 + \mathbf{u}_0$ y la otra por los puntos de la forma $t\mathbf{v}_1 + \mathbf{v}_0$, son **paralelas** si $\mathbf{v}_1$ es múltiplo numérico de $\mathbf{u}_1$. Toda recta $s\mathbf{u}_1 + \mathbf{u}_0$ es paralela a una que pasa por el origen: la recta $s\mathbf{u}_1$.

Ejemplo 1. Para encontrar la representación de la recta en $\mathcal{R}^3$ paralela al vector (1, 1, 1) y que pasa por el punto (−1, 3, 6), formamos todos los múltiplos $t(1, 1, 1)$ para obtener así la recta que pasa por el origen. Entonces, el conjunto de todos los puntos de la forma $t(1, 1, 1) + (-1, 3, 6)$ es una recta que contiene al punto (−1, 3, 6), como se comprueba haciendo $t = 0$.

Para determinar un plano en $\mathcal{R}^3$ es natural partir de dos vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ que *no son colineales* (es decir, tales que ninguno es un múltiplo del otro) y considerar todos los puntos $u\mathbf{x}_1 + v\mathbf{x}_2$, donde u y v son números. La interpretación geométrica de la multiplicación numérica y de la suma de vectores muestra que los puntos $u\mathbf{x}_1 + v\mathbf{x}_2$ forman lo que nos gustaría llamar un plano P_0 que pasa por el origen. Entonces definimos un **plano** P paralelo a P_0 como el conjunto de todos los puntos $u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$, donde $\mathbf{x}_0$ es un vector fijo. La relación entre P y P_0 se representa en la Fig. 9.

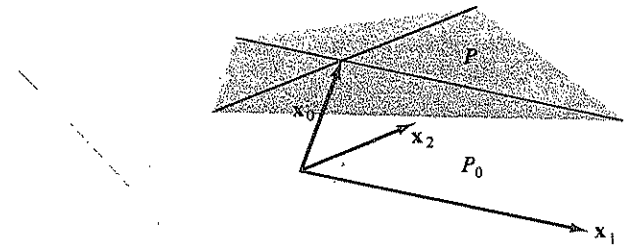


Figura 9

Ejemplo 2. Consideremos el plano en $\mathcal{R}^3$ paralelo a los dos vectores (1, 1, 1) y (2, −1, 3) y que pasa por el punto (0, 1, −1). El conjunto de todas las combinaciones lineales $u(1, 1, 1) + v(2, -1, 3)$ es un plano P_0 que contiene el origen y el conjunto de todos los puntos $u(1, 1, 1) + v(2, -1, 3) + (0, 1, -1)$ es un plano paralelo P. Se comprueba que P contiene al punto (0, 1, −1) haciendo $u = v = 0$.

El conjunto de todas las combinaciones lineales de un conjunto S se llama el **espacio generado** por S, o se dice que S genera al conjunto de todas las combinaciones lineales de S. Por ejemplo, si S está formado por un solo vector $\mathbf{x}$, entonces la recta $\mathcal{L}$ formada por todos los múltiplos numéricos $t\mathbf{x}$, es el espacio generado por S. Si $S = \{\mathbf{x}_1, \mathbf{x}_2\}$, el espacio generado por S es el conjunto P de todas las combinaciones lineales $u\mathbf{x}_1 + v\mathbf{x}_2$. Para que P sea un plano es necesario que $\mathbf{x}_1$ y $\mathbf{x}_2$ no estén en la misma recta. Otra forma de decir esto es que ningún múltiplo de $\mathbf{x}_1$ puede ser un múltiplo de $\mathbf{x}_2$. En forma más general, decimos que un conjunto

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$$

de vectores es **linealmente independiente** si cada vez que

$$c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = 0$$

para ciertos números $c_1, \ldots, c_n$, entonces todos los c deben ser cero. Por otra parte, si es posible encontrar una combinación lineal de vectores que sea igual al vector cero sin que todos los coeficientes sean cero, entonces esos vectores se dicen **linealmente dependientes**.

Ejemplo 3. El conjunto de los tres vectores

$$(1, 2), (-1, 1), (1, -3)$$

en $\mathcal{R}^2$ es linealmente dependiente, ya que la ecuación

$$c_1(1, 2) + c_2(-1, 1) + c_3(1, -3) = 0$$

es equivalente a las ecuaciones

$$c_1 - c_2 + c_3 = 0$$
$$2c_1 + c_2 - 3c_3 = 0.$$

Si c_3 tiene un valor distinto de cero, podemos despejar c_1 y c_2. Por ejemplo, si $c_3 = 1$, entonces

$$c_1 - c_2 = -1$$
$$2c_1 + c_2 = 3,$$

y resolvemos estas ecuaciones. Sumándolas se tiene que $3c_1 = 2$ o $c_1 = \frac{2}{3}$, mientras que restando la segunda de dos veces la primera da $-3c_2 = -5$ o $c_2 = \frac{5}{3}$. Por tanto,

$$(\tfrac{2}{3})(1, 2) + (\tfrac{5}{3})(-1, 1) + (1)(2, -3) = 0,$$

de modo que los vectores son linealmente dependientes. La independencia lineal puede ser comprobada en forma análoga resolviendo una ecuación vectorial y demostrando que solo tiene soluciones cero.

Un conjunto de vectores linealmente independientes que generan un subconjunto $\mathcal{S}$ de un espacio vectorial $\mathcal{V}$ se llama una **base** de $\mathcal{S}$. Una base de $\mathcal{S}$ es útil ya que todo vector $\mathbf{x}$ de $\mathcal{S}$ puede ser representado como una combinación lineal de elementos de la base, con coeficientes determinados unívocamente. Así, si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ es una base de $\mathcal{S}$ y $\mathbf{x}$ está en $\mathcal{S}$, tenemos, por la propiedad de generación, que

$$\mathbf{x} = c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n,$$

para algunas constantes $c_1, \ldots, c_n$. Pero las c están completamente determinadas, ya que si también

$$\mathbf{x} = d_1\mathbf{x}_1 + \ldots + d_n\mathbf{x}_n,$$

sustrayendo una ecuación de la otra, se obtiene

$$0 = (c_1 - d_1)\mathbf{x}_1 + \ldots + (c_n - d_n)\mathbf{x}_n.$$

Por la independencia lineal de los $\mathbf{x}$, se tiene que $c_k - d_k = 0$ para todo k, de modo que $c_k = d_k$.

Ejemplo 4. La base natural de $\mathcal{R}^3$ es el conjunto de vectores

$$\mathbf{e}_1 = (1, 0, 0), \quad \mathbf{e}_2 = (0, 1, 0), \quad \mathbf{e}_3 = (0, 0, 1).$$

Es fácil verificar que estos vectores son linealmente independientes y que ellos generan a $\mathcal{R}^3$. En efecto, si

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + c_3\mathbf{e}_3 = 0,$$

entonces $(c_1, c_2, c_3) = (0, 0, 0)$. Por otra parte, todo vector (x, y, z) en $\mathcal{R}^3$ puede ser escrito en la forma

$$(x, y, z) = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3.$$

EJERCICIOS

1. Para cada par de vectores $\mathbf{u}$, $\mathbf{v}$ en $\mathcal{R}^2$ dados más adelante, dibuje las flechas que representan a $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, $\mathbf{u} - \mathbf{v}$ y $\mathbf{u} + 2\mathbf{v}$
   (a) $\mathbf{u} = (1, 0)$, $\mathbf{v} = (0, 1)$.
   (b) $\mathbf{u} = (-2, 1)$, $\mathbf{v} = (1, 2)$.
   (c) $\mathbf{u} = (-1, 1)$, $\mathbf{v} = (\frac{1}{2}, \frac{1}{2})$.

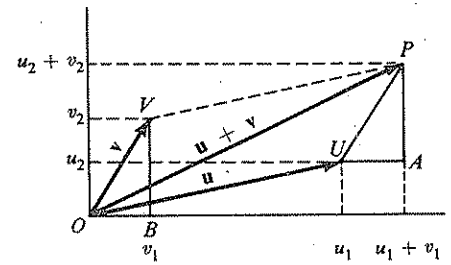
2. Sea $\mathbf{x} = (1, 1)$ e $\mathbf{y} = (0, 1)$. Dibuje las flechas que representan a $t\mathbf{x} + \mathbf{y}$ para los siguientes valores de t: -1, $\frac{1}{2}$, 1, 2.

3. Sea $\mathbf{u} = (2, 1)$ y $\mathbf{v} = (-1, 2)$. Dibuje los vectores que representan $t\mathbf{u} + (1 - t)\mathbf{v}$ para los siguientes valores de t: -1, 0, $\frac{1}{2}$, 1, 2.

4. (a) Sean $\mathbf{u}_1 = (0, 1)$, $\mathbf{v}_1 = (-1, 1)$, $\mathbf{u}_2 = (-3, 2)$, $\mathbf{v}_2 = (2, 1)$. Dibuje las rectas $\mathbf{u}_1 + t\mathbf{v}_1$ y $\mathbf{u}_2 + s\mathbf{v}_2$. Encuentre el vector $\mathbf{w}$ en la intersección de las rectas calculando los valores de t y s para los cuales $\mathbf{w} = \mathbf{u}_1 + t\mathbf{v}_1 = \mathbf{u}_2 + s\mathbf{v}_2$.
   [*Resp.* $\mathbf{w} = (-\frac{5}{3}, \frac{8}{3})$.]
   (b) Sean $\mathbf{u}_1$, $\mathbf{v}_1$ y $\mathbf{u}_2$ como en la parte (a), pero tome $\mathbf{v}_2 = (2, -2)$. (Note que entonces $\mathbf{v}_2$ es un múltiplo numérico de $\mathbf{v}_1$). Dibuje las rectas $\mathbf{u}_1 + t\mathbf{v}_1$ y $\mathbf{u}_2 + s\mathbf{v}_2$. Demuestre algebraicamente que estas rectas no se cortan.

5. Demuestre que las rectas en $\mathcal{R}^n$ representadas por $s\mathbf{u}_1 + \mathbf{u}_0$ y $t\mathbf{v}_1 + \mathbf{v}_0$ son iguales si, y solo si, $\mathbf{v}_1$ y $\mathbf{v}_0 - \mathbf{u}$ son múltiplos numéricos de $\mathbf{u}_1$.

6.



   (a) En la figura los puntos U, V y P son tales que ellos tienen las coordenadas indicadas. Demuestre, usando geometría plana, que los triángulos OVB y UPA son congruentes. Deduzca de esto que OV y UP son paralelos y de igual largo. Entonces $OVPU$ es un paralelogramo, de lo cual se deduce la regla del paralelogramo.
   (b) En las Figs. 7(a) y 7(b), los puntos U, V y W tienen las coordenadas indicadas. demuestre que los triángulos OUP, OVQ y OWR son similares. Por tanto, demuestre que los ángulos UOP, VOQ y WOR son iguales y que los largos OV y OW son proporcionales al largo de OU como se indicara en el texto.

7. No es esencial usar ejes perpendiculares para introducir coordenadas en el plano. Se puede usar el mismo procedimiento si se define la proyección de un punto sobre un eje como la intersección de ese eje con la recta paralela al otro eje que pasa por el punto dado. (Si los ejes son perpendiculares, esto es lo mismo que la definición dada anteriormente. ¿Por qué?

También es posible elegir diferentes unidades de distancia en cada eje. Demuestre que la interpretación geométrica de las operaciones vectoriales en dos dimensiones es la misma en esta situación más general. ¿Como podría extender la definición de proyección al caso de tres dimensiones de modo que se generalice lo anterior?

8. (a) Demuestre que si $\mathbf{u}$ y $\mathbf{v}$ son dos vectores distintos, entonces los vectores $t\mathbf{u} + (1-t)\mathbf{v}$ forman una recta que pasa por los puntos que corresponden a $\mathbf{u}$ y $\mathbf{v}$.
(b) Si $\mathbf{x}_1$ y $\mathbf{x}_2$ son dos vectores de $\mathcal{R}^n$, el conjunto de todos los vectores $t\mathbf{x}_1 + (1-t)\mathbf{x}_2$, donde $0 \leq t \leq 1$ es el segmento que une a $\mathbf{x}_1$ y $\mathbf{x}_2$. Un subconjunto S de $\mathcal{R}^n$ es **convexo** si para todo par de puntos de S, el segmento que los une también está contenido en S. Demuestre que la intersección de cualquier colección de conjuntos convexos es convexo.

9. Represente las siguientes rectas en la forma $t\mathbf{x}_1 + \mathbf{x}_0$, donde t recorre los reales. Dibuje cada recta.

(a) La recta en $\mathcal{R}^3$ paralela a (1, 2, 0) que pasa por el punto (1, 1, 1).
(b) La recta en $\mathcal{R}^2$ que pasa por los puntos (1, 0) y (0, 1).
(c) La recta en $\mathcal{R}^3$ que pasa por los puntos (1, 0, 0) y (0, 0, 1).

10. Represente los planos en $\mathcal{R}^3$ siguientes en la forma $u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$, donde u y v recorren los reales. Dibuje cada plano.

(a) El plano paralelo a los vectores (1, 1, 0) y (0, 1, 1) que pasa por el origen.
(b) El plano paralelo a los vectores $\mathbf{e}_1$ y $\mathbf{e}_2$ que pasa por el punto (0, 0, 1).
(c) El plano que pasa por los puntos (1, 0, 0), (0, 1, 0) y (0, 0, 1).

11. Determine si cada uno de los siguientes conjuntos de vectores son linealmente dependientes o linealmente independientes.

(a) (1, 2), (2, 1).
(b) (1, 2), (−3, −6).
(c) (1, −1), (0, 1), (2, 5).
(d) (1, 2, 0), (−1, 1, 1), (−1, 0, 0).

12. Determine si el primer vector está en el conjunto generado por S. En cada caso dé una interpretación geométrica: el primer punto está (o no está) en el plano (o recta) generada por S.

(a) (1, 1); $S = \{(-2, -2)\}$.
(b) (−1, 1, 3); $S = \{(1, 1, 1), (-1, 0, 1)\}$.
(c) (−1, −2, 1); $S = \{3, 6, -3), (1, 2, -1)\}$.

13. ¿Cuáles de los conjuntos siguientes son bases de $\mathcal{R}^2$ o $\mathcal{R}^3$?

(a) (1, 1), (−1, 1).
(b) (1, 2, 1), (−1, 2, 1), (0, 4, 2).
(c) (1, 1, 1), (1, 1, 0), (1, 0, 0).
(d) (1, 0, 0), (0, 2, 0), (0, 0, 3).

14. (a) Demuestre que dos vectores $\mathbf{x}$ e $\mathbf{y}$ distintos de cero son linealmente dependientes si, y solo si, ellos están en la misma recta por el origen.
(b) Demuestre que tres vectores $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ distintos de cero son linealmente independientes si, y solo si, ellos están en el mismo plano por el origen.

3. Matrices

Un conjunto de ecuaciones tales como

$$y_1 = 2x_1 + 3x_2 - 4x_3$$
$$y_2 = x_1 - x_2 + 2x_3,$$

en las cuales cada y_i está dada como una suma de múltiplos constantes de las x_j define una función (en el ejemplo, de $\mathcal{R}^3$ en $\mathcal{R}^2$) de un tipo particularmente sencillo. Este es un ejemplo de lo que llamaremos función *lineal*. (La definición precisa de lo que significa «lineal» aparece en la sección siguiente.) La comprensión de estas funciones es básica para el estudio de funciones más generales. Las funciones lineales de $\mathcal{R}^1$ en $\mathcal{R}^1$ son tan sencillas, que pueden darse por conocidas al estudiar el cálculo de las funciones de una variable. En dimensiones más altas, las funciones lineales pueden ser más complicadas y el fin principal de este capítulo es desarrollar sus propiedades básicas e introducir la notación y los métodos de cálculo para tratarlas.

En el ejemplo anterior, las x y las y solo sirven para indicar algunos lugares. La función es descrita totalmente por el sistema de coeficientes.

$$\begin{pmatrix} 2 & 3 & -4 \\ 1 & -1 & 2 \end{pmatrix}.$$

Un arreglo rectangular de números, como el indicado anteriormente, se llama una **matriz.** Así, por ejemplo,

$$\begin{pmatrix} 0 & 5 \\ -1 & \frac{1}{2} \\ 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0{,}7 & 3 \\ 0{,}9 & 0 & 2{,}8 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (\tfrac{1}{2}, \tfrac{1}{3}, 0), \quad \begin{pmatrix} 0{,}325 \\ 0{,}007 \\ 0{,}579 \\ 3{,}142 \end{pmatrix}$$

son ejemplos de matrices. Las líneas horizontales de números de una matriz se llaman **filas** y las verticales, **columnas**. El número de filas y de columnas determinan su **forma.** Así, los ejemplos anteriormente mencionados tienen formas 3-por-2, 2-por-3, 2-por-2, 1-por-3 y 4-por-1. Nótese que el número de filas aparece antes que el número de columnas. Las matrices 1-por-n se llaman **vectores fila n-dimensionales** y las matrices n-por-1 se llaman **vectores columnas** n-dimensionales. Una matriz es **cuadrada** si tiene el mismo número de columnas y de filas.

El número en la fila i y columna j de una matriz se llama el **elemento** ij de la matriz. Nótese que el primer índice indica la fila y el segundo la columna. Dos matrices son iguales si, y solo si, ellas tienen la misma forma y los elementos en posiciones correspondientes son iguales.

Usaremos mayúsculas para denotar matrices, y a menudo representaremos sus elementos con la misma letra minúscula, dotada de subíndices adecuados.

Así, por ejemplo, escribimos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \quad \text{o} \quad A = (a_{ij}),$$

Generalmente escribiremos $x_1, x_2, \ldots, x_n$ para los elementos de un vector columna n-dimensional x, en lugar de $x_{11}, x_{21}, \ldots, x_{n1}$.

Las operaciones de suma y multiplicación numérica, definida en la sección anterior para elementos de $\mathcal{R}^n$, puede ser extendida a las matrices. Si A y B tienen la misma forma, entonces su **suma** $A + B$ es, por definición, la matriz con la misma forma tal que su elemento ij es $a_{ij} + b_{ij}$. Por ejemplo,

$$\begin{pmatrix} 2 & 3 \\ -2 & \frac{1}{2} \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 2 & \frac{8}{3} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 0 & \frac{7}{6} \\ 0 & 4 \end{pmatrix}.$$

La suma no está definida entre matrices de distinta forma. Recuerde que no definimos suma entre elementos de $\mathcal{R}^n$ y $\mathcal{R}^m$ si $n \neq m$.

Si A es una matriz y r un número, el **múltiplo numérico** rA se define como la matriz con la misma forma que A y cuyo elemento ij es ra_{ij}. Por ejemplo,

$$2\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

y

$$5\begin{pmatrix} 1 & 0 \\ 3 & 2 \\ -1 & 0 \end{pmatrix} + 3\begin{pmatrix} -2 & 2 \\ -4 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 15 & 10 \\ -5 & 0 \end{pmatrix} + \begin{pmatrix} -6 & 6 \\ -12 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 6 \\ 3 & 10 \\ -5 & 3 \end{pmatrix}.$$

Igual que en el caso de vectores de $\mathcal{R}^n$, escribiremos $-A$ en lugar de $(-1)A$ y $A - B$ en lugar de $A + (-1)B$. Para cada forma hay una matriz cero tal que sus elementos son todos cero. Usaremos el símbolo 0 para denotar cualquier matriz cero; su forma quedará clara del contexto. Es fácil ver que si las matrices X, Y y Z tienen la misma forma y r y s son números cualquiera, entonces las fórmulas 1-7 de la Sec. 1, son satisfechas. (Las demostraciones son iguales al caso en que X, Y y Z están en $\mathcal{R}^n$.) En otras palabras, de acuerdo con la definición de la Sec. 2, para m y n fijos, el conjunto de matrices m-por-n es un espacio vectorial con las operaciones de suma y multiplicación numérica que acabamos de definir.

Otra operación entre matrices es sugerida por la manera en que ellas pueden ser usadas para escribir funciones de $\mathcal{R}^m$ en $\mathcal{R}^n$. Por ejemplo, supongamos que tenemos fórmulas

$$z_1 = 3y_1 - y_2$$
$$z_2 = 5y_1 + 2y_2,$$

y

$$y_1 = 2x_1 + 3x_2 + 4x_3$$
$$y_2 = x_1 - x_2 + 2x_3,$$

que definen funciones de $\mathcal{R}^2$ en $\mathcal{R}^2$ y de $\mathcal{R}^3$ en $\mathcal{R}^3$, respectivamente. Estas funciones son descritas por las matrices A y B, donde

$$A = \begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix}.$$

Si expresamos las z directamente en términos de las x, tenemos

$$z_1 = 3(2x_1 + 3x_2 + 4x_3) - (x_1 - x_2 + 2x_3)$$
$$z_2 = 5(2x_1 + 3x_2 + 4x_3) + 2(x_1 - x_2 + 2x_3).$$

Reordenando términos, tenemos

$$z_1 = (3 \cdot 2 - 1 \cdot 1)x_1 + (3 \cdot 3 - 1 \cdot (-1))x_2 + (3 \cdot 4 - 1 \cdot 2)x_3$$
$$z_2 = (5 \cdot 2 + 2 \cdot 1)x_1 + (5 \cdot 3 + 2 \cdot (-1))x_2 + (5 \cdot 4 + 2 \cdot 2)x_3$$

y vemos que la función que así resulta está descrita por la matriz C tal que

$$C = \begin{pmatrix} 3 \cdot 2 - 1 \cdot 1 & 3 \cdot 3 - 1 \cdot (-1) & 3 \cdot 4 - 1 \cdot 2 \\ 5 \cdot 2 + 2 \cdot 1 & 5 \cdot 3 + 2 \cdot (-1) & 5 \cdot 4 + 2 \cdot 2 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 10 \\ 12 & 13 & 24 \end{pmatrix}.$$

Decimos que C se obtiene de A y B por multiplicación y escribimos $C = AB$. Para ver cómo debe darse la definición general, nótese que el elemento ij de C es la de productos de los elementos en la fila i de A y en la columna j de B. Así, $c_{21} = a_{21}$, $b_{11} + a_{22}b_{21} = 5 \cdot 2 + 2 \cdot 1 = 12$.

Daremos la definición de la multiplicación de matrices en dos etapas. Primero supongamos que

$$A = (a_1, a_2, \ldots, a_k) \quad \text{y} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_k \end{pmatrix}$$

son vectores fila y columna de la misma dimensión. Entonces el producto AB se define como número $a_1b_1 + a_2b_2 + \ldots + a_kb_k$. Ahora bien, sea A una matriz m-por-k y B una matriz k-por-n. (Es importante que el número de columnas de A sea igual al número de filas de B). Entonces, cada fila de A es un vector fila k-dimensional y cada columna de B un vector columna k-dimensional. Definimos el **producto matricial** AB como la matriz m-por-n cuyo elemento ij es el

producto (en el sentido ya definido) de la fila i de A con la columna j de B. El producto AB tiene el mismo número de filas que A y el mismo número de columnas que B. Por ejemplo, en nuestro ejemplo

$$\begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 10 \\ 12 & 13 & 24 \end{pmatrix},$$

el elemento en la segunda fila y tercera columna del resultado se obtiene por el cálculo

$$(5 \quad 2)\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 5 \cdot 4 + 2 \cdot 2 = 24.$$

El lector deberá comprobar que los otros elementos se obtienen de la manera indicada.

$$\begin{pmatrix} * & * & * & * \\ \boxed{*} & * & * & * \\ * & * & * & * \end{pmatrix}\begin{pmatrix} * & * & \boxed{*} & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & \boxed{*} & * \\ * & * & * & * & * \end{pmatrix}$$

La indicación siguiente es una consecuencia obvia de la manera como hemos definido la multiplicación matricial. La expresamos formalmente para darle énfasis y porque nos referiremos a ella en el futuro.

3.1 Teorema

La fila i del producto matricial AB es igual a la fila i de A por B. La columna j de AB es igual a A por la columna j de B.

Hay varias leyes importantes que relacionan la multiplicación matricial con las operaciones de adición y multiplicación numérica. Ellas son válidas para un número t y matrices A, B, C cualesquiera, siempre que las operaciones indicadas estén definidas. (La suma está definida solo entre matrices de la misma forma. La multiplicación está definida solo si el primer factor tiene tantas columnas como filas tiene el segundo.)

1. $(A + B)C = AC + Bc$.
2. $A(B + C) = AB + AC$.
3. $(tA)B = t(AB) = A(tB)$.
4. $A(BC) = (AB)C$.

De acuerdo con la última ley, tiene sentido hablar *del* producto de tres matrices y escribir simplemente ABC, ya que el resultado es independiente de cómo se agrupen los factores. En el hecho, esta ley de asociatividad para productos de tres factores implica que el resultado de una multiplicación de una sucesión finita de matrices es independiente de la manera cómo se agrupen. No todas las leyes que son válidas para la multiplicación de números son válidas para la multiplicación de matrices. En particular, el valor de un producto de matrices depende del orden de los factores y AB es generalmente distinto de BA. Además, también es posible que el producto de dos matrices sea la matriz cero sin que ninguno de los factores sea cero. El Ejercicio 5, al final de esta sección ilustrará estos puntos.

Las leyes expresadas anteriormente se demuestran fácilmente escribiendo lo que significan, usando las definiciones de las operaciones y aplicando las leyes de asociatividad, distributividad y conmutatividad de la aritmética. El número 4 es el más difícil de demostrar y daremos su demostración. Las otras demostraciones son dejadas como ejercicios.

Para demostrar que $A(BC) = (AB)C$, supongamos que A, B, C tienen formas p-por-q, q-por-r y r-por-s, respectivamente. Sea $U = BC$ y $V = AB$. (Entonces U tiene forma q-por-s y V tiene forma p-por-r.) Debemos demostrar que $AU = VC$. El elemento ij de AU es (por definición de multiplicación matricial) igual a $\sum_{k=1}^{q} a_{ik}u_{kj}$. El elemento kj de U es $\sum_{l=1}^{r} b_{kl}c_{lj}$. Por tanto, el elemento ij de AU es $\sum_{k=1}^{q} a_{ik}\left(\sum_{l=1}^{r} b_{kl}c_{lj}\right)$. Del mismo modo, el elemento ij de VC es igual a $\sum_{l=1}^{r} v_{il}c_{lj} = \sum_{l=1}^{r}\left(\sum_{k=1}^{q} a_{ik}b_{kl}\right)c_{lj}$. Estas dos expresiones son iguales a la suma

$$\sum_{\substack{1\le l\le r \\ 1\le k\le q}} a_{ik}b_{kl}c_{lj}$$

y, por tanto, son iguales entre sí. Entonces, los elementos correspondientes de AU y VC son iguales y las matrices son las mismas. Esto termina la demostración.

Una matriz cuadrada de la forma

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{o} \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{o} \quad I = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ . & . & & . & . \\ . & . & & . & . \\ . & . & & . & . \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

que tiene unos en su diagonal principal y ceros en todos los demás lugares se llama la **matriz identidad**. Tiene la propiedad de que

$$IA = A, \quad BI = B$$

para todas las matrices A, B tales que los productos están definidos. Es un elemento identidad para la multiplicación matricial, tal como el número 1 es un elemento identidad para la multiplicación de números. Hay una matriz iden-

tidad n-por-n para cada valor de n, pero al igual que para las matrices cero, casi siempre es claro del contexto cuál debe ser la forma de la matriz identidad considerada.

Si A es una matriz cuadrada y existe una matriz B (de la misma forma) tal que

$$AB = BA = I,$$

entonces decimos que A es **invertible** y que B es el **inverso** de A. Si A es una matriz, existe a lo sumo una matriz B que satisface la condición anteriormente indicada, ya que si $AB_1 = B_1A = I$ y $AB_2 = B_2A = I$, entonces la identidad $(B_1A)B_2 = B_1(AB_2)$ dice que $IB_2 = B_1I$; es decir, que $B_1 = B_2$. Esto justifica hablar *del* inverso de una matriz. Si A es invertible, denotaremos por A^{-1} su inverso. Por ejemplo, es fácil comprobar que

$$\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix}\begin{pmatrix}7 & -2\\ -3 & 1\end{pmatrix} = \begin{pmatrix}7 & -2\\ -3 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

y, de acuerdo con la definición, esto muestra que

$\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix}$ es invertible y que $\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix}^{-1} = \begin{pmatrix}7 & -2\\ -3 & 1\end{pmatrix}$.

Por otra parte, muchas matrices no son invertibles. Ninguna matriz cero puede ser invertible. Otros ejemplos menos obvios serán dados en los ejercicios.

Generalmente, no es fácil determinar si una matriz grande es invertible o no y puede ser muy complicado encontrar su inversa si ella existe. Los determinantes pueden ser usados para obtener una fórmula para el inverso de una matriz (Teorema 8.3) y en la Sec. 1 del Cap. 2 se verá un método más efectivo para calcular inversos. El caso de matrices 2-por-2 y otros casos simples son tratados en los Ejercicios 9 a 13 de esta sección. La regla para encontrar la inversa de una matriz 2-por-2 es

3.2 $$\begin{pmatrix}a & b\\ c & d\end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}, \quad \text{si} \quad ad - bc \neq 0.$$

Así, por ejemplo,

$$\begin{pmatrix}1 & 3\\ -1 & 2\end{pmatrix}^{-1} = \tfrac{1}{5}\begin{pmatrix}2 & -3\\ 1 & 1\end{pmatrix} = \begin{pmatrix}\frac{2}{5} & -\frac{3}{5}\\ \frac{1}{5} & \frac{1}{5}\end{pmatrix}.$$

Es fácil verificar que

$$\begin{pmatrix}1 & 3\\ -1 & 2\end{pmatrix}\begin{pmatrix}\frac{2}{5} & -\frac{3}{5}\\ \frac{1}{5} & \frac{1}{5}\end{pmatrix} = \begin{pmatrix}\frac{2}{5} & -\frac{3}{5}\\ \frac{1}{5} & \frac{1}{5}\end{pmatrix}\begin{pmatrix}1 & 3\\ -1 & 2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

de modo que efectivamente hemos encontrado la inversa.

Es fácil demostrar las propiedades importantes de las matrices invertibles directamente a partir de las definiciones. La primera asegura que no importa cómo se calcula la matriz inversa de A, ya que siempre se obtiene el mismo resultado A^{-1}.

3.3 Teorema

Una matriz A tiene a lo más una inversa.

Demostración. Supongamos que hay dos matrices B y C tales que

$$AB = BA = I \quad \text{y} \quad AC = CA = I$$

Entonces $AB = AC$, ya que ambos productos son iguales a I.

Multiplicando por la izquierda por B se tiene que

$$BAB = BAC.$$

Pero como $BA = I$, sustituyendo se tiene que $IB = IC$, de donde $B = C$, como se quería demostrar.

El teorema que sigue muestra cómo calcular el inverso de un producto de matrices invertibles.

3.4 Teorema

Si $A = A_1A_2, \ldots, A_n$ y $A_1, \ldots, A_n$ son todas invertibles, entonces A es invertible y $A^{-1} = A_n^{-1}A_{n-1}^{-1} \ldots A_2^{-1}A_1^{-1}$.

Demostración. En el producto $(A_n^{-1} \ldots A_2^{-1}A_1^{-1})(A_1A_2 \ldots A_n)$, el producto $A_1^{-1}A_1$ es I y, por tanto, se puede eliminar. Entonces A_2^{-1} y A_2 se cancelan. Siguiendo de esta manera obtenemos I como resultado final. Del mismo modo demostramos que $(A_1 \ldots A_n)(A_n^{-1} \ldots A_1) = I$, con lo cual queda demostrado que los dos productos son inversos uno del otro.

Además de la identidad hay otras matrices para las cuales es conveniente tener nombres y notaciones especiales. Una matriz A es **diagonal** si los elementos que no están en la «diagonal principal» son todos cero; es decir, si $a_{ij} = 0$ cuando $i \neq j$. La notación **diag** $(t_1, t_2, \ldots, t_n)$ es conveniente para la matriz diagonal n-por-n que tiene los elementos $t_1, t_2, \ldots, t_n$ en la diagonal. Por ejemplo, **diag** $(2, 0, -1, 3)$ es la notación para la matriz

$$\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 3\end{pmatrix}.$$

Los Ejercicios 7 y 8 muestran que las operaciones matriciales son muy sencillas en el caso de matrices diagonales.

EJERCICIOS

1. Dadas las matrices

$$A = \begin{pmatrix} 1 & 3 \\ -4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -2 & 1 \\ -1 & 3 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & 0 \\ 2 & 3 & -1 \end{pmatrix},$$

$$D = \begin{pmatrix} 2 & -4 \\ 0 & 0 \\ 3 & 3 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 0 & 3 \end{pmatrix},$$

determinar cuáles de las siguientes expresiones están definidas y compútelas.

(a) $2B - 3G$ (f) $CD + 3DB$
(b) AB (g) $2AB - 5G$
(c) BA (h) $2GC - AAB$
(d) BD (i) CDC
(e) DB (j) DCD

[*Resp.* (b) $\begin{pmatrix} -3 & 7 & 1 \\ -2 & 14 & -4 \end{pmatrix}$.]

2. Demuestre que para cualquier matriz A y matrices cero de formas adecuadas,

$$A0 = 0 \quad \text{y} \quad 0A = 0.$$

Si A es m-por-n, para matrices cero de qué forma está definido $A0$? ¿Para qué formas está definido $0A$? ¿Cuáles son las formas de los productos?

3. Para las matrices A, B, C, D del Ejercicio 1, determine qué formas deberán tener X e Y para que las ecuaciones siguientes tengan sentido. (En algunos casos no hay forma posible; en otros varias son posibles.)

(a) $AX = B + Y$. (d) $CX + DY = 0$
(b) $(D + 2X)YC = 0$. (e) $AX = YC$.
(c) $AX = YD$. (f) $AX = CY$.

[*Resp.* (d) X si 3-por-n, Y si 2-por-n.]

4. Demuestre la ley distributiva

$$A(B + C) = AB + AC$$

para la multiplicación matricial.

5. Sean $U = \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix}$, $V = \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix}$. Calcule UV y VU. ¿Son iguales? ¿Es posible que el producto de dos matrices sea cero sin que ninguno de los factores sea cero?

6. Sea $X = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $P = \begin{pmatrix} 0 & -1 \\ 4 & 3 \\ 2 & 0 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -4 & -1 \\ -1 & 2 & 0 \end{pmatrix}$.

Sea $D = \mathbf{diag}\,(1, 2, 3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Calcule DX, DP y DQ.

7. Demuestre que, como lo ilustra el Ejercicio 6, el producto DR, donde D es la matriz diagonal $\mathbf{diag}\,(d_1, \ldots, d_n)$ y R es una matriz cualquiera con n filas, se obtiene multiplicando la fila i por d_i, para todo i. Supongamos que S tiene n columnas. ¿Cómo se puede escribir el producto SD? (El cálculo de QD para las matrices del Ejercicio 6 deberá sugerir la regla general).

8. Usando las matrices B, C y G del Ejercicio 1, calcule Be_1, Ce_1, Ge_1, Be_3, Ce_3, Ge_3. Demuestre la regla general que dice que para cualquier matriz M y vector columna $\mathbf{e}_j$ de dimensiones apropiadas, el producto $M\mathbf{e}_j$ es la columna j de M.

9. Calcule el producto $\mathbf{diag}\,(a_1, \ldots, a_n)\ \mathbf{diag}\,(b_1, \ldots, b_n)$. ¿Cuándo es este producto la matriz identidad? Demuestre que $\mathbf{diag}\,(a_1, \ldots, a_n)$ tiene un inverso siempre que ninguno de los números a_i sea cero.

10. (a) Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz 2-por-2 arbitraria y sea B dada por la Ec. 3.2. Calcule AB y BA y demuestre que si $ad \neq bc$, entonces B es el inverso de A (de modo que A es invertible).

(b) Use la Ec. 3.2 para tratar de encontrar las inversas de las siguientes matrices 2-por-2.

$$\begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix}.$$

¿Qué falla en el último caso?

11. Demuestre que si existe una matriz $X \neq 0$ tal que $AX = 0$, entonces A no puede ser invertible. [*Indicación*: Suponga $B = A^{-1}$ y considere $(BA)X = B(AX)$]. Use este resultado para demostrar que $\mathbf{diag}\,(a_1, \ldots, a_n)$ no es invertible si alguno de los a_i es cero.

12. Demuestre que si $ad = bc$, la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ no es invertible.

13. Si A es una matriz cuadrada, ella puede ser multiplicada consigo misma y podemos definir $A^2 = AA$, $A^3 = AAA = A^2A$, $A^n = AA \ldots A$ (n factores). Estas potencias tienen la misma forma que A. Encuentre A^2 y A^3 si

(a) $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 2 & 1 & -1 \end{pmatrix}$.

[*Resp.* (a) $A^2 = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$.]

(Note que 0 es el único número cuyo cubo es 0. La parte (b) de este ejemplo ilustra otra diferencia entre la aritmética de los números y la de las matrices.)

14. La ecuación numérica $a^2 = 1$ tiene las soluciones únicas $a = 1$ y $a = -1$.

(a) Demuestre que si $A = I$ o $-I$, entonces $A^2 = I$, donde I es una matriz identidad de cualquier dimensión.

(b) Demuestre que $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ si $a^2 + bc = 1$, de modo que la ecuación $A^2 = I$ tiene infinitas soluciones distintas en el conjunto de las matrices 2-por-2.

(c) Demuestre que toda matriz A tal que $A^2 = I$, es I, $-I$ o una de las matrices descritas en (b).

4. Funciones lineales

El producto de una matriz m-por-n y de un vector columna (matriz n-por-1) es un vector columna m-dimensional. Obviamente, a una n-pla en $\mathcal{R}^n$ corresponde un vector columna n-dimensional único y viceversa. De aquí en adelante, a menudo consideraremos los elementos de $\mathcal{R}^n$ como vectores columnas. Con esta convención, la multiplicación por una matriz m-por-n dada, define una función de $\mathcal{R}^n$ en $\mathcal{R}^m$. En efecto, una ecuación matricial como

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -4 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

es equivalente al conjunto de ecuaciones numéricas

$$y_1 = 2x_1 + 3x_2 - 4x_3$$
$$y_1 = x_1 - x_2 + 2x_3,$$

como se puede comprobar escribiendo el resultado de la multiplicación matricial. Por tanto, la función vectorial definida por una matriz consiste en multiplicar los elementos del dominio por la matriz.

Las funciones definidas por multiplicación matricial pueden ser caracterizadas por propiedades muy sencillas. Nótese que las definiciones que daremos son aplicables a cualquier par de espacios vectoriales, a pesar de que, por ahora, nos referiremos solo a los espacios $\mathcal{R}^n$.

Una función f cuyo dominio es un espacio vectorial $\mathcal{U}$ y cuya imagen es un subconjunto de un espacio vectorial $\mathcal{W}$ es una función **lineal** si las ecuaciones

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$$
$$f(r\mathbf{x}) = rf(\mathbf{x})$$

son válidas para todos los vectores $\mathbf{x}$, y en $\mathcal{U}$ y todos los números r.

4.1 Teorema

Una función de $\mathcal{R}^n$ en $\mathcal{R}^m$ es lineal si, y solo si, ella es la multiplicación por alguna matriz m-por-n.

Demostración. Primero demostraremos que si f está definida por $f(\mathbf{x}) = A\mathbf{x}$ para una matriz fija A, entonces f es lineal. Debemos demostrar que $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ y $f(r\mathbf{x}) = rf(\mathbf{x})$ para todo $\mathbf{x}$, $\mathbf{y}$ y r. Por la definición de f, esto es lo mismo que $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ y $A(r\mathbf{x}) = rA\mathbf{x}$, relaciones que son verdaderas por las propiedades 2 y 3 de la multiplicación matricial.

La demostración de la afirmación conversa es un poco más complicada. Supongamos dada una función lineal g de $\mathcal{R}^n$ en $\mathcal{R}^m$. Debemos encontrar una matriz m-por-n, A tal que $g(\mathbf{x}) = A\mathbf{x}$ para todo vector columna $\mathbf{x}$ de $\mathcal{R}^n$.

Sea $\mathbf{e}_j$ el vector columna de $\mathcal{R}^n$ cuyo elemento j es 1 y cuyos elementos distintos de j son 0. Sea A la matriz cuya columna j es el vector m-dimensional $g(\mathbf{e}_j$ para $j = 1, 2, \ldots, n)$. (Así, A tiene m filas y n columnas, como se desea.) Cualquier vector columna $\mathbf{x}$ con elementos $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ puede ser expresado en la forma $x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. Como g es lineal,

$$g(\mathbf{x}) = g(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n)$$
$$= x_1 g(\mathbf{e}_1) + \ldots + x_n g(\mathbf{e}_n).$$

Por definición de A tenemos

$$g(\mathbf{x}) = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2n}x_n \\ \cdot \\ \cdot \\ \cdot \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}$$

Finalmente, por definición de multiplicación matricial,

$$g(\mathbf{x}) = A\mathbf{x}.$$

La demostración dada nos enseña cómo construir la matriz que corresponde a una función lineal. Esta construcción es importante y la damos como un teorema para darle mayor énfasis.

4.2 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ una función lineal y sea A la matriz cuya columna j es $f(\mathbf{e}_j)$. Entonces $f(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ en $\mathcal{R}^n$. Es fácil recordar este resultado considerando, por ejemplo, el caso de una función f de $\mathcal{R}^2$ en $\mathcal{R}^2$. Entonces la matriz A para la cual $f(\mathbf{x}) = A\mathbf{x}$ es 2-por-2 y sus columnas son el resultado de aplicar A sucesivamente a $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}.$$

Ejemplo 1. Consideremos la función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^2$ descrita, en términos de la representación geométrica de $\mathcal{R}^2$, como la rotación en 30° en el sentido contrario al de los punteros del reloj. Toda rotación deja el origen fijo, preserva distancias y transforma figuras tales como paralelogramos, en figuras congruentes. A partir de estas propiedades y de las construcciones geométricas que se pueden hacer la suma y multiplicación numérica de vectores (Sec. 2), se puede demostrar que las rotaciones son funciones lineales. La Fig. 10(a) muestra los vectores $\mathbf{e}_1$, $\mathbf{e}_2$, $f(\mathbf{e}_1)$ y $f(\mathbf{e}_2)$. Usando trigonometría, veremos que

$$f(\mathbf{e}_1) = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} \quad \text{si} \quad f(\mathbf{e}_2) = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Por el Teorema 4.2,

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2}x_1 - \frac{1}{2}x_2 \\ \frac{1}{2}x_1 + \frac{\sqrt{3}}{2}x_2 \end{pmatrix}$$

para todo vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ en $\mathcal{R}^2$.

Ejemplo 2. Definimos $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^2$ como la función lineal que multiplica las distancias horizontales por 3 y las verticales por 2. Entonces $f(\mathbf{e}_1) = 3\mathbf{e}_1$ y $f(\mathbf{e}_2) = 2\mathbf{e}_2$ y f tiene la matriz diagonal

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

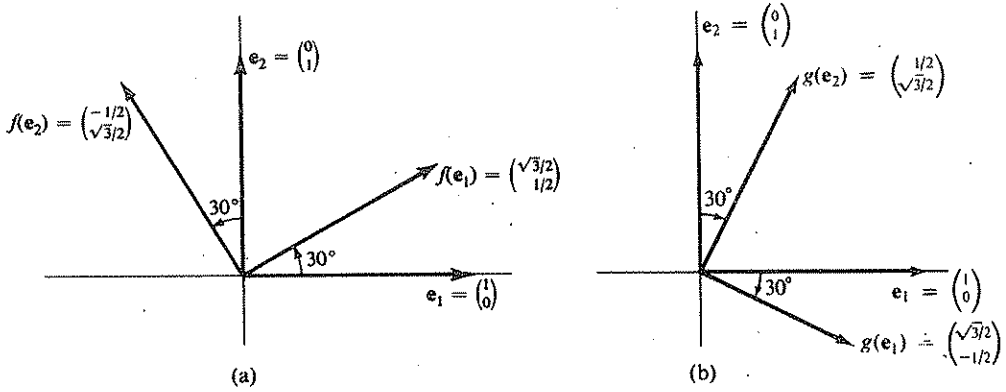


Figura 10

El efecto geométrico de f está ilustrado en la Fig. 11, en la cual C es el círculo unitario $x_1^2 + x_2^2 = 1$ y $f(C)$ es la imagen de C por f. Si $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ es la imagen por f de $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ entonces $x_1 = \frac{1}{3}u_1$ y $x_2 = \frac{1}{2}u_2$. Por tanto, si $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ está en $f(C)$, $\frac{u_1^2}{9} + \frac{u_2^2}{4} = 1$ la que puede ser reconocida como la ecuación de una elipse cuyo semieje mayor es 3 y cuyo semieje menor es 2.

El Teorema 4.2 dice que si f es una función lineal de $\mathcal{R}^n$ en $\mathcal{R}^m$, entonces la columna j de la matriz de f es $f(\mathbf{e}_j)$.

Como todo vector $\mathbf{x}$ en $\mathcal{R}^n$ puede ser expresado como una combinación lineal

$$\mathbf{x} = c_1\mathbf{e}_1 + \ldots + c_n\mathbf{e}_n,$$

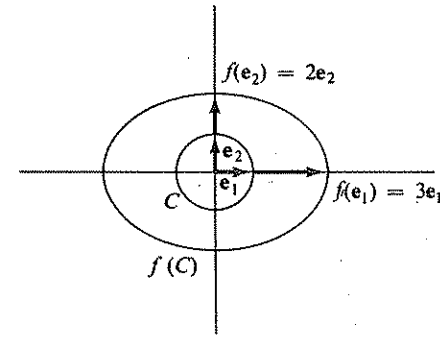


Figura 11

podemos usar la linealidad de f para escribir

$$f(\mathbf{x}) = c_1 f(\mathbf{e}_1) + \ldots + c_n f(\mathbf{e}_n).$$

Como $\mathbf{x}$ es un vector arbitrario en el dominio de f, esta última ecuación expresa un vector arbitrario en la imagen de f como una combinación lineal de los vectores $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$. Pero estos vectores son precisamente las columnas de la matriz de f, de modo que tenemos que

4.3 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ una función lineal con matriz A. Entonces las columnas de A, consideradas como vectores en $\mathcal{R}^m$, generan la imagen de f.

Ejemplo 3. Si P es un plano por el origen en $\mathcal{R}^3$, entonces P está forma-

do por todas las combinaciones lineales de dos vectores $\mathbf{y}_1$ e $\mathbf{y}_2$ en $\mathcal{R}^3$. La función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^3$ definida por

$$f\begin{pmatrix} u \\ v \end{pmatrix} = u\mathbf{y}_1 + v\mathbf{y}_2$$

es lineal (¿por qué?) y el Teorema 4.2 dice que la matriz de f tiene como columnas los dos vectores

$$f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbf{y}_1 \quad \text{y} \quad f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{y}_2.$$

Por ejemplo, si

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{y} \quad \mathbf{y}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

entonces

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix},$$

y tenemos que la imagen de f está generada por las columnas de la matriz 3-por-2.

Ejemplo 4. Si L es una recta formada por todos los puntos de la forma $t\mathbf{u}_1 + \mathbf{u}_0$ con $\mathbf{u}_1 \neq 0$ y f es una función lineal, entonces

$$f(t\mathbf{u}_1 + \mathbf{u}_0) = tf(\mathbf{u}_1) + f(\mathbf{u}_0)$$
$$= t\mathbf{v}_1 + \mathbf{v}_0.$$

Por tanto, a no ser que $f(\mathbf{u}_1) = 0$, la imagen $f(L)$ también es una recta. Si $f(\mathbf{u}_1) = 0$, entonces, por supuesto, $f(L)$ es único punto $f(\mathbf{u}_0)$. La ecuación $f(\mathbf{u}_1) = 0$ no puede ser válida para un vector $\mathbf{u}_1$ distinto de cero si la función lineal es inyectiva de acuerdo con la definición siguiente:

Una función f es **inyectiva** si cada punto de la imagen de f corresponde a un solo punto de su dominio. En otras palabras, f es inyectiva si la ecuación $f(\mathbf{x}) + f(\mathbf{y})$ y $f(r\mathbf{x}) = rf(\mathbf{x})$ para todo $\mathbf{x}$, $\mathbf{y}$ y r. Por definición de f, esto es definidas por $f(x) = x$ y $g(x) = x^2$, es obvio que f es inyectiva. Pero g no lo es, ya que $g(x) = g(-x)$ para todo número x. En cuanto a funciones *lineales*, tenemos el criterio siguiente:

4.4 Teorema

Una función lineal es inyectiva si, y solo si, $f(\mathbf{x}) = 0$ implica $\mathbf{x} = 0$.

Demostración. Primero supongamos que $f(\mathbf{x}) = 0$ implica $\mathbf{x} = 0$. Si $f(\mathbf{x}_1) = f(\mathbf{x}_2)$, por linealidad tenemos que $f(\mathbf{x}_1 - \mathbf{x}_2) = 0$. Pero entonces, por hipótesis, $\mathbf{x}_1 - \mathbf{x}_2 = 0$, de modo que $\mathbf{x}_1 = \mathbf{x}_2$. Por tanto, f es inyectiva. Inversamente, supongamos que $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ implica $\mathbf{x}_1 = \mathbf{x}_2$. Entonces como para una función lineal, $f(0) = 0$, la ecuación $f(\mathbf{x}) = 0$ puede ser escrita como $f(\mathbf{x}) = f(0)$. Pero entonces, por hipótesis, $\mathbf{x} = 0$.

Ejemplo 5. La rotación en 30° en el sentido contrario al de los punteros del reloj considerada en el Ejemplo 1, es claramente inyectiva. En efecto, un solo vector es llevado a cualquier otro mediante una rotación. Lo que es lo mismo, el vector cero es el único vector llevado al vector cero. En términos algebraicos, se ha demostrado en el Ejemplo 1 que esta rotación puede ser representada por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2}x - \frac{1}{2}y \\ \frac{1}{2}x + \frac{\sqrt{3}}{2}y \end{pmatrix}$$

Entonces el Teorema 4.4 muestra que f es inyectiva si, y solo si, las ecuaciones

$$\frac{\sqrt{3}}{2}x - \frac{1}{2}y = 0$$
$$\frac{1}{2}x + \frac{\sqrt{3}}{2}y = 0$$

tienen la solución única $(x, y) = (0, 0)$. Esto puede ser comprobado resolviendo las ecuaciones.

Si f y g son dos funciones (no necesariamente lineales) tales que el espacio imagen de f es igual al espacio dominio de g, definimos la **composición** $g \circ f$ como la función obtenida aplicando primero f y después g. Más explícitamente, si $\mathbf{x}$ está en el dominio de f y $f(\mathbf{x})$ está en el dominio de g, entonces $g \circ f(\mathbf{x})$ es, por definición, $g(f(\mathbf{x}))$. Si $f(\mathbf{x})$ o $g(f(\mathbf{x}))$ no está definida, entonces $g \circ f(\mathbf{x})$ no está definida.

La composición de funciones lineales es la razón para la definición de multiplicación matricial. Al introducir el concepto de multiplicación matricial, consideramos una función de $\mathcal{R}^2$ en $\mathcal{R}^2$ dada por

$$z_1 = 3y_1 - y_2$$
$$z_2 = 5y_1 + 2y_2$$

y otra de $\mathcal{R}^3$ en $\mathcal{R}^2$ definida por

$$y_1 = 2x_1 + 3x_2 + 4x_3$$
$$y_2 = x_1 - x_2 + 2x_3.$$

y después observamos que la composición de las dos funciones estaba dada por las fórmulas

$$z_1 = 5x_1 + 10x_2 + 10x_3$$
$$z_2 = 12x_1 + 13x_2 + 24x_3.$$

La definición de multiplicación matricial fue diseñada de tal manera que la matriz $\begin{pmatrix} 5 & 10 & 10 \\ 12 & 13 & 24 \end{pmatrix}$ de la función compuesta fuera el producto de las matrices

$$\begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix}$$

de las funciones originales. El teorema siguiente afirma el resultado importante que la composición de funciones lineales está dada siempre por la multiplicación matricial.

4.5 Teorema

Sean f y g funciones lineales tales que $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$, $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^p$ están dadas por las matrices A, B. Es decir, $f(\mathbf{x}) = A\mathbf{x}$ y $g(\mathbf{y}) = B\mathbf{y}$ para todo $\mathbf{x}$ en $\mathcal{R}^n$ e $\mathbf{y}$ en $\mathcal{R}^m$. Entonces la composición $g \circ f$ está dada por la matriz BA, de modo que $g \circ f(\mathbf{x}) = BA\mathbf{x}$ para todo $\mathbf{x}$ en $\mathcal{R}^n$.

Demostración. Nótese que A es una matriz m-por-n y B una matriz p-por-m, de modo que el producto BA está bien definido y es de la forma adecuada; es decir, p-por-n. Por definición de $g \circ f$, $g \circ f(\mathbf{x}) = g(f(\mathbf{x})) = B(A\mathbf{x})$ para todo $\mathbf{x}$ en $\mathcal{R}^n$. Por la asociatividad de la multiplicación matricial, esto es igual a $(BA)\mathbf{x}$, como se quería demostrar.

Ejemplo 6. Consideremos la función $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^2$ definida como $f \circ f \circ f$, donde f es la función del Ejemplo 1. Por el Teorema 4.5, $g(\mathbf{x}) = B\mathbf{x}$, donde

$$B = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Por tanto, $g(\mathbf{e}_1) = B\mathbf{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{e}_2$ y $g(\mathbf{e}_2) = B\mathbf{e}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\mathbf{e}_1$. Como

lo muestra la Fig. 12, geométricamente, g es una rotación de 90° en el sentido opuesto al de los punteros del reloj, lo cual es, por supuesto, el resultado de tres rotaciones de 30°.

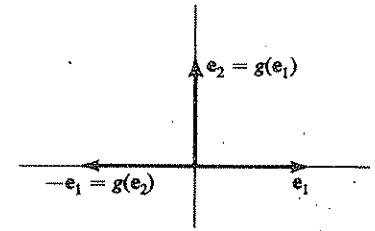


Figura 12

Ejemplo 7. La función identidad $\mathcal{R}^n \xrightarrow{h} \mathcal{R}^n$ tal que $h(\mathbf{x}) = \mathbf{x}$, es una función lineal. Su matriz es la matriz identidad n-por-n, I. Dos funciones $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ y $\mathcal{R}^n \xrightarrow{g} \mathcal{R}^n$ se llaman **inversas** una de la otra si $f \circ g$ y $g \circ f$ son la función identidad. Por ejemplo, si $n = 2$, supongamos que f es la rotación en 30° en el sentido contrario al de los punteros del reloj, discutida en el Ejemplo 1, y supongamos que g es la rotación en 30° en el sentido de los punteros del reloj. Las funciones f y g son evidentemente inversas una de la otra. Usando la Fig. 12, se obtiene que las matrices de f y g son

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{1} \end{pmatrix},$$

respectivamente. Multiplicando estas matrices en los dos órdenes posibles, se obtiene la matriz identidad 2-por-2. Por tanto, las matrices de f y g son inversas una de la otra. Más generalmente, dos funciones lineales de $\mathcal{R}^n$ en $\mathcal{R}^n$ son inversas una de la otra si, y solo si, sus matrices son inversas una de la otra.

Al tratar las funciones de una variable, es habitual decir que una función de la forma $f(x) = ax + b$ es una función lineal. Esta terminología no es consistente con nuestra definición de linealidad. Por ejemplo, si $f(x) = 2x + 1$, entonces $f(x + y) = 2(x + y) + 1$, mientras que $f(x) + f(y) = 2x + 2y + 2$, de modo que esta función no satisface la condición $f(x + y) = f(x) + f(y)$, como se requiere para la linealidad. Como nuestra definición de función lineal entre espacios vectoriales es la estándar, no consideraremos a las funciones del tipo $2x + 1$ como funciones lineales. Sin embargo, las funciones de este tipo son importantes y existe un nombre estándar para ellas. Se las llama funciones *afines* y están caracterizadas por tener la forma de una función lineal (como $2x$) más una constante (como 1). A continuación, damos la definición formal.

Si $\mathbf{b}$ es un elemento de un espacio vectorial $\mathcal{U}$, entonces la función $t_{\mathbf{b}}$ de $\mathcal{U}$ en sí mismo definida por $t_{\mathbf{b}}(\mathbf{x}) = \mathbf{x} + \mathbf{b}$ para todo $\mathbf{x}$ en $\mathcal{U}$ se llama la **traslación** de $\mathcal{U}$ inducida por $\mathbf{b}$.

Ejemplo 8. Sea $\mathbf{b}$ el vector $(1, 0)$ en $\mathcal{R}^2$. Entonces $t_{\mathbf{b}}$ envía el vector (x, y) en el vector $(x + 1, y)$. En lenguaje geométrico, cada punto del plano es movido en una unidad hacia la derecha.

Una función de un vector $\mathcal{U}$ en un vector $\mathcal{W}$ es una función **afín** si ella es la composición de una función lineal de $\mathcal{U}$ a $\mathcal{W}$ con una traslación de $\mathcal{W}$.

Ejemplo 9. Sea f de $\mathcal{R}^2$ en $\mathcal{R}^2$ dada por la matriz $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ y sea $\mathbf{b}$ el vector $(0, 2)$. La función afín $t_{\mathbf{b}} \circ f$ envía el vector (x, y) en el vector $(x, -y + 2)$, geométricamente, esto consiste en una reflexión en el eje x seguida de un movimiento hacia arriba, de dos unidades. Es fácil ver que todos los puntos de la recta $y = 1$ son dejados fijos por esta función y que la función puede ser descrita como una reflexión en la recta $y = 1$.

El teorema siguiente es una consecuencia inmediata de las definiciones y el Teorema 4.1.

4.6 Teorema

Una función f de $\mathcal{R}^n$ en $\mathcal{R}^m$ es afín si, y solo si, $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ para alguna matriz m-por-n, A y un vector columna m-dimensional $\mathbf{b}$.

Es fácil encontrar la matriz A y el vector $\mathbf{b}$ necesarios para la descripción de una función afín. Si $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, entonces $\mathbf{b} = f(0)$. Entonces la función $g(\mathbf{x}) = f(\mathbf{x}) - f(0)$ es lineal y su matriz puede ser encontrada usando el Teorema 4.2.

Ejemplo 10. Supongamos que la función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^2$ está definida en términos de la representación geométrica estándar como una rotación de 90° alrededor del centro $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbf{e}_1$, y en sentido opuesto al de los punteros del reloj. Vemos que $f(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $f(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (puesto que el centro de rotación permanece fijo) y $f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Introduciendo la función lineal $g(\mathbf{x}) = f(\mathbf{x}) - f(0)$, tenemos que $g(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, y $g(\mathbf{e}_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Por tanto, $g(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x}$ para todo $\mathbf{x}$ y

$$f(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Como comprobación, calculemos

$$f\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

lo cual está de acuerdo con la descripción geométrica de f. Véase Fig. 13.

Las funciones lineal y afín a menudo son llamadas **transformaciones** lineal y afín, aunque usemos más a menudo, en este libro, el término función.

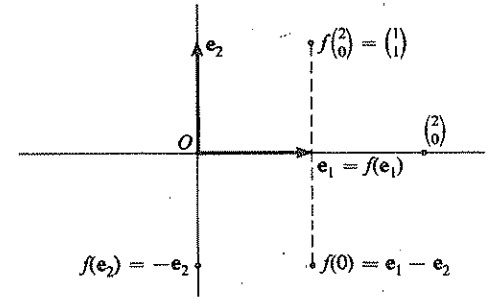


Figura 13

EJERCICIOS

1. (a) Demuestre que las dos condiciones para una función lineal dadas en el texto son equivalentes a la condición única

$$f(r\mathbf{x} + s\mathbf{y}) = rf(\mathbf{x}) + sf(\mathbf{y})$$

para todo $\mathbf{x}$, $\mathbf{y}$ en el dominio de f y números r, s.

(b) Demuestre por inducción en n que si f es lineal, entonces

$$f(r_1\mathbf{x}_1 + \ldots + r_{n-1}\mathbf{x}_{n-1} + r_n\mathbf{x}_n) = r_1 f(\mathbf{x}_1) + \ldots + r_{n-1} f(\mathbf{x}_{n-1}) + r_n f(\mathbf{x}_n)$$

para números $r_1, \ldots, r_n$ y vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ cualesquiera.

2. Demuestre que $f(x) = ax + b$ es una función lineal (en la terminología de este libro) si, y solo si, $b = 0$.

3. Sean f y g las funciones lineales de $\mathcal{R}^2$ en $\mathcal{R}^2$ tales que

(a) $f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

(b) $g\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $g\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Encuentre, en cada caso, la matriz que representa a la función.

4. Demuestre que una rotación de $\mathcal{R}^2$ en un ángulo α en el sentido opuesto al de los punteros del reloj está dada por la matriz $\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$. Sea β otro ángulo. Calcule $R_\alpha R_\beta$. La composición de una rotación en un ángulo α seguida de otra en un ángulo β es una rotación en un ángulo $\alpha + \beta$. ¿Cuál es la relación entre $R_\alpha R_\beta$ y $R_{\alpha+\beta}$? ¿Cuál es el inverso de R_α?

5. (a) Demuestre que la matriz $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ define una función lineal de $\mathcal{R}^2$ en $\mathcal{R}^2$ que corresponde geométricamente a una reflexión en una recta que pasa por el origen y tiene una inclinación de 45° con respecto a la horizontal, en el sentido opuesto al de los punteros del reloj.

(b) ¿Qué matriz corresponde a la reflexión en la recta que pasa por el origen y que tiene una inclinación de 135° con respecto a la horizontal, en el sentido opuesto al de los punteros del reloj?

(c) Calcule el producto de las matrices de (a) y (b) e interprete geométricamente el resultado.

[*Resp.* (b) $\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$.]

6. (a) Encuentre una matriz 2-por-2 M_α que corresponda a la reflexión en una recta que pasa por el origen y que forma un ángulo α con la horizontal. Compruebe sus resultados comparando los casos $\alpha = 45°$ y $\alpha = 135°$ con el Ejercicio 5. ¿Cuál es M_α^2?

(b) Sea β otro ángulo. Calcule $M_\alpha M_\beta$. Demuestre que representa una rotación, identifique el ángulo de rotación. ¿Cuándo se cumple $M_\alpha \circ M_\beta = M_\beta \circ M_\alpha$?

7. (a) Demuestre que

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{y} \quad V = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

representan rotaciones en 90° de $\mathcal{R}^3$ alrededor del eje x_1 y x_2 respectivamente. Encuentre la matriz W que representa la rotación en 90° alrededor del eje x_3. Encuentre también U^{-1} y V^{-1} (que representan rotaciones en el sentido opuesto).

(b) Calcule UVU^{-1} y VUV^{-1} e interprete geométricamente los resultados. (Posiblemente será conveniente que manipule un modelo.)

8. Demuestre que una función f de un espacio vectorial en otro es afín si, y solo si,

$$f(r\mathbf{x} + (1 - r)\mathbf{y}) = rf(\mathbf{x}) + (1 - r)f(\mathbf{y})$$

para todos los números r y todos los vectores $\mathbf{x}$, $\mathbf{y}$ en el dominio de f. [*Indicación*: considere la función $g(\mathbf{x}) = f(\mathbf{x}) - f(0)$.]

9. Sea f una función lineal de un espacio vectorial $\mathcal{V}$ en un espacio vectorial $\mathcal{W}$. Demuestre que si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ son vectores linealmente independientes en $\mathcal{V}$ y f es inyectiva, entonces los vectores $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)$ son linealmente independientes en $\mathcal{W}$. [*Sugerencia*: Si f es inyectiva y $f(\mathbf{x}) = 0$, entonces $\mathbf{x} = 0$].

10. (a) Demuestre que si f es una función lineal de $\mathcal{R}^3$ en $\mathcal{R}^3$ y f es inyectiva, entonces la imagen $f(L)$ de una recta L es una recta.

(b) Dé un ejemplo que muestre que si la función lineal de la parte (a) no es inyectiva, entonces la imagen de una recta puede ser un punto.

(c) Demuestre que si L_1 y L_2 son rectas paralelas y f es una función lineal inyectiva entonces $f(L_1)$ y $f(L_2)$ son rectas paralelas.

11. Demuestre que la composición de dos funciones afines de un espacio en sí mismo, es también afín. Supongamos que $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, $g(\mathbf{x}) = C\mathbf{x} + \mathbf{d}$. Supongamos que $(f \circ g)(\mathbf{x}) = P\mathbf{x} + \mathbf{q}$. Exprese P y $\mathbf{q}$ en función de A, $\mathbf{b}$, C y $\mathbf{d}$. ¿Cuándo se tiene que $f \circ g = g \circ f$?

12. Sea $\mathbf{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{b} = -\mathbf{a} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, y sea $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^2$ una rotación de 90° en el sentido opuesto al de los punteros del reloj con centro en el origen. Encuentre la matriz que corresponde a la función lineal g y calcule la función afín $f = t_{\mathbf{a}} \circ g \circ t_{\mathbf{b}}$, donde $t_{\mathbf{a}}$ y $t_{\mathbf{b}}$ son las translaciones inducidas por $\mathbf{a}$ y $\mathbf{b}$.

13. Demuestre que una función afín es inyectiva si, y solo si, $A(\mathbf{x}) = A(0)$ implica $\mathbf{x} = 0$.

14. ¿Cuáles de las funciones lineales siguientes son inyectivas?

(a) $f(x_1, x_2) = (x_1 + 2x_2, 2x_1 + x_2)$.

(b) $g(x_1, x_2, x_3) = (x_2 - x_3, x_3 - x_1, x_1 - x_2)$.

5. Producto punto

Para poder aplicar cabalmente los conceptos vectoriales a la geometría euclidiana, debemos disponer de algún método para introducir conceptos tales como largo, ángulo y perpendicularidad. En esta sección demostraremos que todos estos conceptos pueden ser definidos si introducimos una nueva operación entre vectores. Si $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_n)$ son vectores en $\mathcal{R}^n$, definimos el **producto punto** o **producto interno** de $\mathbf{x}$ e $\mathbf{y}$ como el número

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + \ldots + x_ny_n.$$

Es fácil verificar (véase Ejercicio 4) que el producto punto de vectores en $\mathcal{R}^n$ satisface las propiedades siguientes.

5.1 **Positividad:** $\mathbf{x} \cdot \mathbf{x} > 0$ excepto por $0 \cdot 0 = 0$

Conmutatividad: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

Distributividad: $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$.

Asociatividad: $(r\mathbf{x}) \cdot \mathbf{y} = r(\mathbf{x} \cdot \mathbf{y})$.

Debido a la conmutatividad del producto punto, se tiene de inmediato que la distributividad y asociatividad son válidas también para el segundo vector, es decir,

$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$$

y

$$\mathbf{x} \cdot (r\mathbf{y}) = r(\mathbf{x} \cdot \mathbf{y}).$$

Consideremos primero el largo de un vector $\mathbf{x}$ de $\mathcal{R}^3$. Si $\mathbf{x} = (x_1, x_2, x_3)$, consideramos el largo del vector $\mathbf{x}$ como la distancia desde el origen al punto de coordenadas (x_1, x_2, x_3). En $\mathcal{R}^3$ el teorema de Pitágoras nos da una fórmula para esta distancia (véase Fig. 15). Si denotamos por $|\mathbf{x}|$ el **largo** del vector $\mathbf{x}$, tenemos

$$|\mathbf{x}|^2 = x_1^2 + x_2^2 + x_3^2. \tag{1}$$

Vemos así que el **largo** de un vector puede ser expresado en términos del producto punto como $\sqrt{\mathbf{x} \cdot \mathbf{x}} = |\mathbf{x}|$. Nótese que usamos para el largo de un vector el mismo símbolo que para el valor absoluto de un número. En efecto, si pensamos en un número como un vector con una coordenada, su largo es su valor absoluto. En $\mathcal{R}^n$ definimos el **largo** de un vector de la misma manera que en $\mathcal{R}^3$: $|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$.

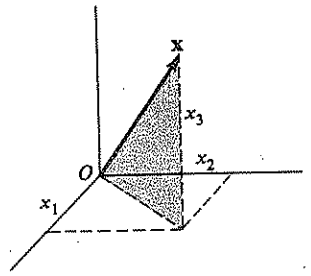


Figura 14

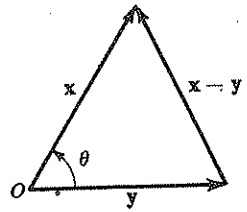


Figura 15

A continuación quisiéramos expresar el ángulo entre dos vectores distintos de cero de $\mathcal{R}^n$. La convención habitual es la de tomar este ángulo θ en el intervalo $0 \leq \theta \leq \pi$ (véase Ejercicio 1). La solución a este problema está dada por el siguiente teorema.

5.2 Teorema

Si θ es el ángulo entre $\mathbf{x}$ e $\mathbf{y}$, entonces

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|\,|\mathbf{y}|}.$$

Demostración. Apliquemos la ley del coseno al triángulo de la Fig. 15. Esta dice que

$$|\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\,|\mathbf{x}|\,|\mathbf{y}| \cos\theta,$$

lo cual puede ser reescrito en la forma

$$(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - 2\,|\mathbf{x}|\,|\mathbf{y}| \cos\theta.$$

Expandiendo el miembro de la izquierda, tenemos

$$\mathbf{x} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - 2\,|\mathbf{x}|\,|\mathbf{y}| \cos\theta.$$

De donde,

$$2\mathbf{x} \cdot \mathbf{y} = 2\,|\mathbf{x}|\,|\mathbf{y}| \cos\theta,$$

y el teorema se obtiene dividiendo por $2\,|\mathbf{x}|\,|\mathbf{y}|$.

De aquí se deduce que $\mathbf{x} \cdot \mathbf{y}$ en valor absoluto es a lo más $|\mathbf{x}|\,|\mathbf{y}|$, es decir, $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|$. Esto se conoce como la desigualdad de Cauchy-Schwarz, que será demostrado más generalmente en el Teorema 5.4.

Ejemplo 1. ¿Cuál es el ángulo θ entre $\mathbf{x} = (1, 3)$ e $\mathbf{y} = (-1, 1)$? Fácilmente calculamos

$$\mathbf{x} \cdot \mathbf{x} = 1^2 + 3^2 = 10,$$

$$\mathbf{y} \cdot \mathbf{y} = (-1)^2 + 1^2 = 2,$$

y

$$\mathbf{x} \cdot \mathbf{y} = -1 + 3 = 2.$$

de donde

$$|\mathbf{x}| = \sqrt{10}, \quad |\mathbf{y}| = \sqrt{2}, \quad \text{y} \quad \cos\theta = \frac{1}{\sqrt{5}}.$$

Consultando una tabla de funciones trigonométricas, encontramos $\theta = 1{,}1$ radianes aproximadamente (cerca de 63°).

El teorema nos proporciona además una manera sencilla de comprobar la perpendicularidad de dos vectores. Ellos son perpendiculares si, y solo si, $\theta = \pi/2$ y, por tanto, $\cos\theta = 0$. Por tanto, la condición para la perpendicularidad es simplemente

$$\mathbf{x} \cdot \mathbf{y} = 0. \tag{2}$$

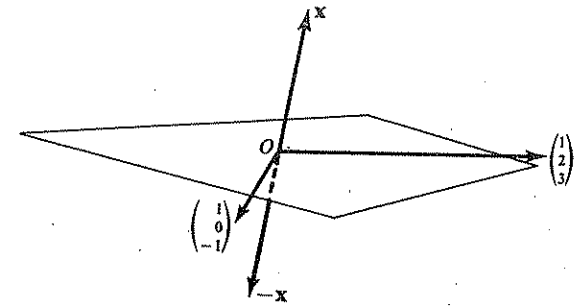


Figura 16

Ejemplo 2. Encontramos un vector $\mathbf{x}$ de largo 2, perpendicular a $(1, 2, 3)$ y a $(1, 0, -1)$. Por consideraciones geométricas, vemos que habrá dos soluciones, ya que si $\mathbf{x}$ es una solución, también lo será $-\mathbf{x}$. (Véase Fig. 16.) Tenemos que escribir tres condiciones, dos para los requisitos de perpendicularidad, usando (2), y una tercera condición que asegure que el largo es 2.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0, \\ x_1 \qquad - x_3 &= 0, \\ x_1^2 + x_2^2 + x_3^2 &= 4. \end{aligned}$$

Estas ecuaciones tienen el par de soluciones $\mathbf{x} = \pm(\sqrt{\frac{2}{3}}, -\sqrt{\frac{8}{3}}, \sqrt{\frac{2}{3}})$.

Si $\mathbf{n}$ es un vector unitario, es decir, de largo 1, entonces el producto punto $\mathbf{n} \cdot \mathbf{x}$ se llama la **coordenada** de $\mathbf{x}$ en la dirección de $\mathbf{n}$. La interpretación geométrica de $\mathbf{n} \cdot \mathbf{x}$ está dada en la Fig. 17. Puesto que $\cos\theta = (\mathbf{n} \cdot \mathbf{x})/|\mathbf{x}|$, se tiene que $\mathbf{n} \cdot \mathbf{x}$ es el largo de la proyección perpendicular en la recta que contiene a $\mathbf{n}$, o su negativo. El vector $(\mathbf{n} \cdot \mathbf{x})\mathbf{n}$ se llama la **componente** de $\mathbf{x}$ en la dirección de $\mathbf{n}$ y se denota a veces por $\mathbf{x}_n$.

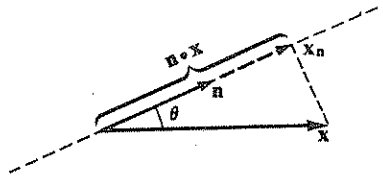


Figura 17

Ejemplo 3. Supongamos que el vector $(1, -1, 2)$ representa una fuerza $\mathbf{F}$ que actúa sobre un punto en el sentido de que la dirección de la fuerza es la misma que del vector y que la magnitud de la fuerza es igual al largo del vector; es decir, $\sqrt{6}$. En una interpretación de este tipo es habitual dibujar el vector como una flecha con su origen en el punto de coordenadas $(1, -1, 2)$ en lugar del origen. La Fig. 18 ilustra una posibilidad.

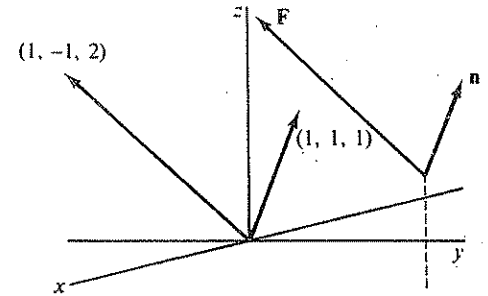


Figura 18

La componente de la fuerza $\mathbf{F}$ en la dirección de un vector unitario $\mathbf{n}$ es interpretada entonces como la fuerza resultante al aplicar $\mathbf{F}$ a un punto que solo puede moverse en una recta paralela a $\mathbf{n}$. Por tanto, para encontrar la componente de $\mathbf{F} = (1, -1, 2)$ en la dirección de $(1, 1, 1)$, primero encontramos el vector unitario en esa dirección. Como $|(1, 1, 1)| = \sqrt{3}$, el vector unitario buscado es $\mathbf{n} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. La componente de $\mathbf{F}$ en la dirección de $\mathbf{n}$ es, entonces,

$$\begin{aligned} (\mathbf{F} \cdot \mathbf{n})\mathbf{n} &= (1, -1, 2) \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\mathbf{n} \\ &= \frac{2}{\sqrt{3}}\mathbf{n} = \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right). \end{aligned}$$

Un espacio vectorial en el cual se ha definido un producto con las propiedades 5.1 se llama un **espacio con producto interno**. Así, $\mathcal{R}^n$ es un espacio con producto interno y otros ejemplos son dados en los Ejercicios 8 y 9. Solo en el Cap. 2, Sec. 4 y en el Cap. 5, Sec. 5, usaremos productos internos en espacios distintos de $\mathcal{R}^n$. La **longitud** o **norma** de un vector puede ser definida siempre mediante un producto interno como

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

Entonces la longitud tiene las propiedades.

5.3 Positividad: $|\mathbf{x}| > 0$ excepto que $|0| = 0$.

Homogeneidad: $|r\mathbf{x}| = |r|\,|\mathbf{x}|$.

Desigualdad triangular: $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$.

Las demostraciones de las dos primeras son sencillas y se dejan al lector. La demostración de la tercera es más difícil y será hecha más adelante. Su significación geométrica se ve fácilmente de la Fig. 19.

El hecho de que la longitud haya sido definida usando un producto interno conduce a propiedades que no son derivables de las propiedades dadas en 5.3. Primero demostraremos

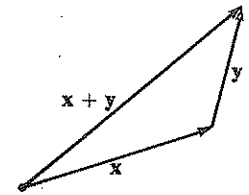


Figura 19

5.4 Desigualdad de Cauchy-Schwarz

$$|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|.$$

Demostración. Primero, supongamos que $\mathbf{x}$ e $\mathbf{y}$ son vectores unitarios; es decir, que $|\mathbf{x}| = |\mathbf{y}| = 1$. Entonces

$$0 \leq |\mathbf{x} - \mathbf{y}|^2 = (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})$$
$$= |\mathbf{x}|^2 - 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2 = 2 - 2\mathbf{x} \cdot \mathbf{y},$$

o

$$\mathbf{x} \cdot \mathbf{y} \leq 1.$$

Si suponemos que $\mathbf{x}$ e $\mathbf{y}$ son distintos de cero (ya que la desigualdad es claramente verdadera si uno de ellos es cero), podemos remplazar $\mathbf{x}$ e $\mathbf{y}$ por los vectores unitarios $\mathbf{x}/|\mathbf{x}|$ e $\mathbf{y}/|\mathbf{y}|$ obteniendo

$$\mathbf{x} \cdot \mathbf{y} \leq |\mathbf{x}|\,|\mathbf{y}|.$$

Ahora, remplazando $\mathbf{x}$ por $-\mathbf{x}$, obtenemos

$$-\mathbf{x} \cdot \mathbf{y} \leq |\mathbf{x}|\,|\mathbf{y}|.$$

Estas dos últimas desigualdades implican la desigualdad de Cauchy-Schwarz.

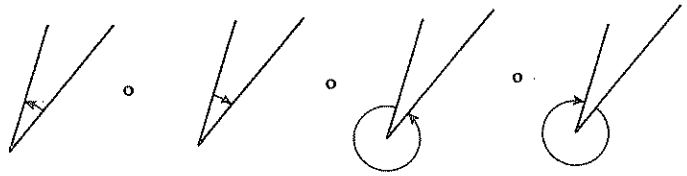
Nótese que si la desigualdad de Cauchy-Schwarz se escribe en la forma

$$\frac{|\mathbf{x} \cdot \mathbf{y}|}{|\mathbf{x}|\,|\mathbf{y}|} \leq 1,$$

entonces existe un ángulo θ tal que

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|\,|\mathbf{y}|}.$$

La definición del coseno del ángulo θ entre $\mathbf{x}$ e $\mathbf{y}$ es más satisfactoria que la definición de θ, ya que no tenemos que preocuparnos de si el ángulo es



Usando la desigualdad de Cauchy-Schwarz, es fácil dar la demostración de la desigualdad triangular. En efecto, de

$$|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|,$$

obtenemos

$$|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = |\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2$$
$$\leq |\mathbf{x}|^2 + 2\,|\mathbf{x}|\,|\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2,$$

de donde

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|.$$

Dos vectores $\mathbf{x}$ e $\mathbf{y}$ que son perpendiculares con respecto a un producto interno se llaman a veces **ortogonales.** Un conjunto S de vectores unitarios que son mutuamente ortogonales se llama un **conjunto ortonormal.** Esta idea tiene una aplicación útil en el problema de encontrar la inversa de algunas matrices. Una matriz cuadrada cuyas columnas consideradas como vectores en $\mathcal{R}^n$ forman un conjunto ortonormal con respecto al producto punto euclidiano, se llama una **matriz ortonormal.** Es muy fácil calcular el inverso de una tal matriz. Supongamos que las columnas de

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot \\ a_{21} & a_{22} & \cdot & \cdot \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \end{pmatrix}$$

forman un conjunto ortonormal. Entonces, por ejemplo, $a_{11}^2 + a_{21}^2 + \ldots + = 1$, $a_{12}^2 + a_{22}^2 + \ldots + = 1$ y $a_{11}a_{12} + a_{21}a_{22} + \ldots + = 0$. Formamos la matriz A^t, llamada la **transpuesta** de A, reflejando A en su diagonal principal. Por tanto,

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \cdot & \cdot \\ a_{12} & a_{22} & \cdot & \cdot \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \end{pmatrix}.$$

El hecho de que las columnas de A son ortonormales, muestra que $A^t = I$. Veremos más adelante (Teorema 8.3) que esto implica que también $AA^t = I$. Un argumento similar remplazando columnas ortonormales por filas ortonormales demuestra que $AA^t = I$ y entonces (por el Teorema 8.3) que $A^tA = I$. Por tanto, hemos demostrado

5.5 Teorema

Toda matriz ortonormal A es invertible y $A^{-1} = A^t$.

Ejemplo 4. Es fácil verificar que la matriz

$$A = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

tiene columnas (y filas) ortonormales. La matriz transpuesta es

$$A^t = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix},$$

de modo que $A^{-1} = A^t$.

La segunda de las dos matrices consideradas anteriormente, se obtiene de la primera reflejando A en su diagonal principal. Cuando dos matrices están relacionadas de esta manera, decimos que una es la **transpuesta** de la otra. Denotamos por A^t la transpuesta de la matriz A. Por ejemplo, si

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

entonces

$$A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Es claro que $(A^t)^t = A$ para toda matriz A.

EJERCICIOS

1. Demuestre que los vectores básicos naturales satisfacen $\mathbf{e}_i \cdot \mathbf{e}_i = 1$ y $\mathbf{e}_i \cdot \mathbf{e}_j = .$ si $i \neq j$.

2. (a) Encuentre el ángulo que forman los vectores (1, 1, 1) y (1, 0, 1).
(b) Encuentre un vector de largo 1 perpendicular a los dos vectores de la parte (a).

3. Encuentre la distancia entre (1, 2) y (0, 5).

4. Demuestre que el producto punto satisface las propiedades dadas en 5.1.

5. Demuestre las propiedades positividad y asociatividad del largo dadas en 5.5.

6. Encuentre la *coordenada* del vector $\mathbf{x} = (1, -1, 2)$:
(a) en la dirección de $\mathbf{n} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ y
(b) en la dirección del vector no unitario (1, 1, 3),
(c) ¿Cuál es la *componente* de $\mathbf{x}$ en la dirección de $\mathbf{n}$?

7. (a) Demuestre que si $\mathbf{x}$ es un vector cualquiera de $\mathcal{R}^n$ y $\mathbf{n}$ es un vector unitario, entonces $\mathbf{x} = \mathbf{y} + \mathbf{z}$, donde $\mathbf{y}$ es un múltiplo de $\mathbf{n}$ y $\mathbf{z}$ es perpendicular a $\mathbf{n}$ [*Indicación*: Tome como $\mathbf{y}$ la componente de $\mathbf{x}$ en la dirección de $\mathbf{n}$.]
(b) Demuestre que los vectores $\mathbf{y}$ y $\mathbf{z}$ de la parte (a) están determinados unívocamente. El vector $\mathbf{z}$ determinado así, se llama la **componente de $\mathbf{x}$ perpendicular a $\mathbf{n}$.**

8. (a) Considere el espacio vectorial $\mathcal{C}[0, 1]$ de todas las funciones reales continuas definidas en el intervalo $0 \leq x \leq 1$. [La suma de f y g está definida por $(f+g)(x) + g(x)$

y la multiplicación numérica rf por $(rf)(x) = rf(x)$.] Demuestre que el producto $\langle f, g \rangle$ definido por

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$$

es un producto interno en $\mathcal{C}[0, 1]$.
(b) La norma de f se define por $|f| = \langle f, f \rangle^{1/2}$. ¿Cuál es la norma de $f(x) = x$?

9. (a) Demuestre que el producto de $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$ definido por

$$\mathbf{x} * \mathbf{y} = x_1y_1 + 2x_2y_2$$

es un producto interno.
(b) Con el largo definido por $|\mathbf{x}|_* = (\mathbf{x} * \mathbf{x})^{1/2}$, dibuje el conjunto de los puntos que satisfacen $|\mathbf{x}|_* = 1$.

10. Demuestre que si $|\mathbf{x}| = |\mathbf{y}|$ en un espacio con producto interno, entonces $\mathbf{x} + \mathbf{y}$ es perpendicular a $\mathbf{x} - \mathbf{y}$.

11. Demuestre que la matriz

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

es ortogonal y encuentre su inversa.

12. Deduzca la desigualdad

$$|\mathbf{x}| - |\mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|$$

de la desigualdad triangular y demuestre que

$$||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|.$$

13. (a) Demuestre que si A es una matriz ortogonal 2-por-2, entonces $|A\mathbf{x}| = |\mathbf{x}|$ para todo $\mathbf{x}$ en $\mathcal{R}^2$.
(b) Demuestre el resultado de la parte (a) para matrices n-por-n.
(c) Demuestre que si $|A\mathbf{x}| = |\mathbf{x}|$ para todo $\mathbf{x}$, entonces $(A\mathbf{x} \cdot A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ para todo $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$.

14. Una función lineal $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ se llama a veces un **funcional lineal.** Demuestre que si f es un funcional lineal definido en $\mathcal{R}^n$, entonces existe un vector fijo $\mathbf{y}$ en $\mathcal{R}^n$ tal que $f(\mathbf{x}) = \mathbf{y} \cdot \mathbf{x}$ para todo $\mathbf{x}$ en $\mathcal{R}^n$. [*Indicación*: ¿Cuál es la matriz de f?]

15. Sean

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -5 & 2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -3 & -1 \\ 1 & 0 \\ 2 & -2 \end{pmatrix}.$$

Calcule AB, BA, A^tB^t y B^tA^t.

16. (a) Demuestre que si A y B son matrices cualesquiera tales que AB está definido, entonces B^tA^t también está definido y $B^tA^t = (AB)^t$.
(b) Demuestre que si A es invertible entonces A^t también lo es y $(A^t)^{-1} = (A^{-1})^t$.

6. Geometría euclideana

En esta sección desarrollaremos algunas propiedades y fórmulas sobre las rectas y planos en $\mathcal{R}^2$ y $\mathcal{R}^3$. Algunas de estas ideas reaparecerán en la Sec. 1 del Cap. 2.

Recordemos que si $\mathbf{x}_1$ es un vector distinto de cero en $\mathcal{R}^2$ o $\mathcal{R}^3$, entonces el conjunto de todos los múltiplos numéricos $t\mathbf{x}_1$ es una recta L_0 que pasa por el origen y que toda recta paralela a L_0 está formada por todos los puntos $t\mathbf{x}_1 + \mathbf{x}_0$, donde $\mathbf{x}_0$ es algún vector fijo. Otra manera de decir esto es que una recta es la imagen de una función de la forma $f(t) = t\mathbf{x}_1 + \mathbf{x}_0$, donde $\mathbf{x}_1 \neq 0$.

Ejemplo 1. Para determinar la recta L que pasa por los dos puntos distintos $\mathbf{a}_1$ y $\mathbf{a}_2$, hagamos $\mathbf{x}_1 = \mathbf{a}_2 - \mathbf{a}_1$. Entonces todos los múltiplos $t(\mathbf{a}_2 - \mathbf{a}_1)$ forman una recta paralela a L y los puntos $t(\mathbf{a}_2 - \mathbf{a}_1) + \mathbf{a}_1$ forman la recta L. Por ejemplo, si $t = 0$, entonces obtenemos $\mathbf{a}_1$ y cuando $t = 1$, obtenemos $\mathbf{a}_2$. Véase la Fig. 20. También podemos escribir los puntos de L en la forma $t\mathbf{a}_2 + (1 - t)\mathbf{a}_1$.

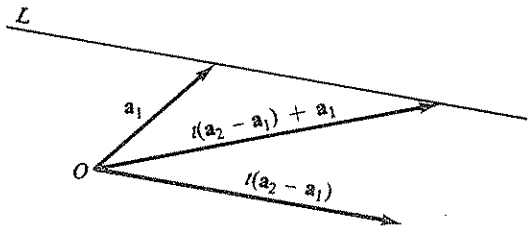


Figura 20

Para determinar un plano en $\mathcal{R}^3$, tomamos dos vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ que no sean colineales (es decir, tales que ninguno de los dos es un múltiplo del otro) y consideremos todos los puntos $u\mathbf{x}_1 + v\mathbf{x}_2$ donde u y v son números. Estos puntos forman un plano P_0 que pasa por el origen. Por tanto, un plano paralelo estará formado por todos los puntos $u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$, donde $\mathbf{x}_0$ es algún vector fijo. Podemos reformular esta definición diciendo que un plano es la imagen de una función $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$, donde $g(u, v) = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$ y los vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ no están en una misma recta.

Ejemplo 2. Para determinar el plano que pasa por los tres puntos $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$ es necesario que ellos no estén en una recta. (En ese caso, los vectores

$\mathbf{a}_3 - \mathbf{a}_1$ y $\mathbf{a}_2 - \mathbf{a}_1$ no son colineales puesto que si $\mathbf{a}_3 - \mathbf{a}_1 = t(\mathbf{a}_2 - \mathbf{a}_1)$, entonces $\mathbf{a}_3 = t\mathbf{a}_2 + (1 - t)\mathbf{a}_1$, de modo que $\mathbf{a}_3$ estaría en la recta determinada por $\mathbf{a}_2$ y $\mathbf{a}_1$.) Ahora tomemos $\mathbf{x}_1 = \mathbf{a}_3 - \mathbf{a}_1$ y $\mathbf{x}_2 = \mathbf{a}_2 - \mathbf{a}_1$. El plano buscado está formado por los puntos

$$\mathbf{w} = u(\mathbf{a}_3 - \mathbf{a}_1) + v(\mathbf{a}_2 - \mathbf{a}_1) + \mathbf{a}_1.$$

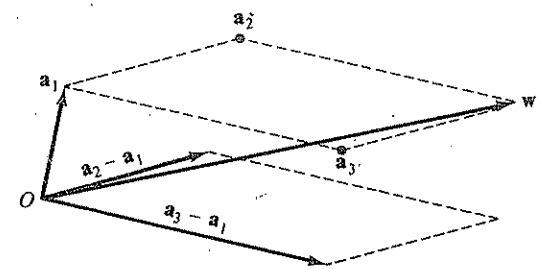


Figura 21

Para obtener $\mathbf{a}_1$, tomamos $u = v = 0$. Para obtener $\mathbf{a}_2$, tomamos $u = 0$, $v = 1$. Para obtener $\mathbf{a}_3$, tomamos $v = 1$, $v = 0$. Véase Fig. 21.

Usando el producto punto se obtiene otra manera de describir un plano P. Supongamos que $\mathbf{x}_0$ es un punto arbitrario de P y que $\mathbf{p}$ es un vector distinto de cero perpendicular a $\mathbf{x}_1$ y a $\mathbf{x}_2$ (de modo que $\mathbf{p} \cdot \mathbf{x}_1 = \mathbf{p} \cdot \mathbf{x}_2 = 0$), donde $\mathbf{x}_1$ y $\mathbf{x}_2$ son vectores paralelos a P. Entonces la ecuación

$$\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \tag{1}$$

es satisfecha por todos los puntos $\mathbf{x} = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$ en P puesto que $\mathbf{p} \cdot (u\mathbf{x}_1 + v\mathbf{x}_2) = u\mathbf{p} \cdot \mathbf{x}_1 + v\mathbf{p} \cdot \mathbf{x}_2 = 0$. También es fácil comprobar que $\mathbf{x}$ puede ser despejado de la Ec. (1), obteniéndose $\mathbf{x} = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$. Dejamos los detalles al lector. La Fig. 22 muestra la relación entre los vectores. Para encontrar un vector $\mathbf{p}$ perpendicular a los vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ es conveniente usar el **producto cruz** de $\mathbf{x}_1$ y $\mathbf{x}_2$, el cual está definido por $\mathbf{x}_1 = (u_1, u_2, u_3)$ y $\mathbf{x}_2 = (v_1, v_2, v_3)$ por

$$\mathbf{p} = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1). \tag{2}$$

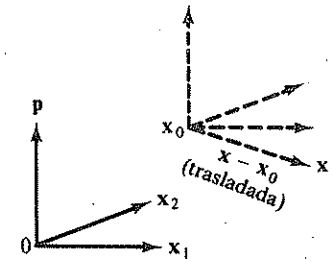


Figura 22

La comprobación de que $\mathbf{p} \cdot \mathbf{x}_1 = 0$ y $\mathbf{p} \cdot \mathbf{x}_2 = 0$ es de rutina. Por ejemplo,

$$\mathbf{p} \cdot \mathbf{x}_1 = u_1u_2v_3 - u_1u_3v_2 + u_2u_3v_1 + u_2u_1v_3 + u_3u_1v_2 - u_3u_2v_1 = 0.$$

El producto cruz será considerado en más detalle en la próxima sección. Por ahora, solo lo usaremos para encontrar vectores perpendiculares a dos vectores dados. Véase Ejercicio 13.

Ejemplo 3. Supongamos que tenemos un plano paralelo a los vectores $\mathbf{x}_1 = (1, 2, -3)$ y $\mathbf{x}_2 = (2, 0, 1)$ y que contienen al punto $\mathbf{x}_0 = (1, 1, 1)$. Un vector $\mathbf{p}$ perpendicular a $\mathbf{x}_1$ y $\mathbf{x}_2$ se obtiene mediante su producto cruz:

$$\begin{aligned}\mathbf{p} &= \big((2)(1) - (0)(-3), (2)(-3) - (1)(1), (1)(0) - (2)(2)\big)\\ &= (2, -7, -4).\end{aligned}$$

Haciendo $\mathbf{x} = (x, y, z)$, pedimos que $\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$; es decir, $(2, -7, -4) \cdot (x - 1, y - 1, z - 1) = 0$. De acuerdo con la definición del producto punto, esta última ecuación es

$$2(x - 1) - 7(y - 1) - 4(z - 1) = 0$$

o

$$2x - 7y - 4z + 9 = 0. \tag{3}$$

En otras palabras, el plano está formado por todos los puntos (x, y, z) que satisfacen la ecuación.

La manera más simple de dibujar un plano que satisface una ecuación como (1) es elegir tres puntos con coordenadas simples que la satisfagan, por ejemplo, $(0, 0, \frac{9}{4})$, $(0, \frac{9}{7}, 0)$ y $(-\frac{9}{2}, 0, 0)$. Entonces se ubican estos puntos con respecto a los ejes de coordenadas y entonces se puede dibujar el plano usando estos tres puntos. En nuestro ejemplo hemos elegido a propósito puntos en los ejes de coordenadas. Véase Fig. 23.

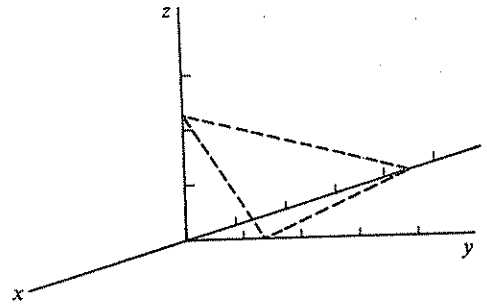


Figura 23

Si $\mathbf{p}$ es un vector distinto de cero en $\mathcal{R}^2$ y $\mathbf{x}_0$ es un punto de $\mathcal{R}^2$, podemos determinar la recta perpendicular a $\mathbf{p}$ que contiene a $\mathbf{x}_0$ mediante la ecuación

$\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$. Si $\mathbf{p} = (p_1, p_2)$ y $\mathbf{x}_0 = (x_0, y_0)$, la ecuación se transforma en

$$(p_1, p_2) \cdot (x - x_0, y - y_0) = 0$$

o

$$p_1(x - x_0) + p_2(y - y_0) = 0.$$

Esta es una de las muchas formas de la ecuación de una recta en el plano x-y. La pendiente de esta recta es claramente $-p_2/p_1$.

La representación de un plano por la ecuación $\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$ no es única, ya que se puede remplazar a $\mathbf{p}$ por cualquier múltiplo no cero de él, sin cambiar el conjunto de los $\mathbf{x}$ que satisfacen la ecuación. Sin embargo, a veces es conveniente normalizar la ecuación del plano exigiendo que $\mathbf{p}$ sea un vector unitario. La ecuación normalizada es entonces $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)$, donde $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$. También podemos escribir $\mathbf{n} \cdot \mathbf{x} = c$, donde $c = \mathbf{n} \cdot \mathbf{x}_0$.

6.1 Teorema

Si $\mathbf{n} \cdot \mathbf{x} = c$ es la ecuación normalizada del plano P e $\mathbf{y}$ es un punto cualquiera, entonces la distancia de $\mathbf{y}$ a P es $c - \mathbf{n} \cdot \mathbf{y}$.

Demostración. Por definición, la distancia debe ser medida a lo largo de una recta perpendicular a P que pase por $\mathbf{y}$. Esta recta puede ser representada por $t\mathbf{n} + \mathbf{y}$ y su intersección con P ocurrirá para algún $t = t_0$. La distancia buscada es, entonces, $|(t_0\mathbf{n} + \mathbf{y}) - \mathbf{y}| = |t_0\mathbf{n}|$ que es, simplemente, el valor absoluto de t_0, ya que $\mathbf{n}$ tiene largo 1. Como $t_0\mathbf{n} + \mathbf{y}$ está en P, tenemos que $\mathbf{n} \cdot (t_0\mathbf{n} + \mathbf{y}) = c$. Pero $\mathbf{n} \cdot \mathbf{n} = 1$, de donde $t_0 + \mathbf{n} \cdot \mathbf{y} = c$, de donde se deduce el teorema.

Ejemplo 4. Encontrar la distancia de $(1, 1, 1)$ al plano $(3, 0, -4) \cdot \mathbf{x} = -3$. La ecuación normalizada del plano es $(\frac{3}{5}, 0, -\frac{4}{5}) \cdot \mathbf{x} = -\frac{3}{5}$. Entonces

$$\begin{aligned}c - \mathbf{n} \cdot \mathbf{y} &= -\tfrac{3}{5} - (\tfrac{3}{5}, 0, -\tfrac{4}{5}) \cdot (1, 1, 1)\\ &= -\tfrac{2}{3}.\end{aligned}$$

Por tanto, la distancia es $\frac{2}{5}$. Nótese que la ecuación del plano podría haber sido escrita en la forma $3x - 4z = -3$ y, en su forma normalizada, $(\frac{3}{5})x - (\frac{4}{5})z = -\frac{3}{5}$.

EJERCICIOS

1. Encuentre una fórmula de la forma $f(t) = t\mathbf{x}_1 + t\mathbf{x}_0$ para la recta paralela a $\mathbf{x}_1$ que contiene a $\mathbf{x}$. En cada caso determine si el punto $\mathbf{a}$ está en la recta, es decir, si $\mathbf{a}$ está en la imagen de f.

   (a) $\mathbf{x}_0 = (1, 1)$, $\mathbf{x}_1 = (-2, 3)$; $\mathbf{a} = (-3, 7)$.
   (b) $\mathbf{x}_0 = (0, 0, 1)$, $\mathbf{x}_1 = (1, 1, 1)$; $\mathbf{a} = (5, 5, 4)$.
   (c) $\mathbf{x}_0 = (1, 0, 2)$, $\mathbf{x}_1 = (2, 0, 3)$; $\mathbf{a} = (3, 0, 5)$.

2. Encuentre una fórmula de la forma $g(u, v) = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$ para el plano paralelo

a $\mathbf{x}_1$ y $\mathbf{x}_2$ que contiene a $\mathbf{x}_0$. En cada caso determine si el punto $\mathbf{a}$ está en el plano, es decir, si $\mathbf{a}$ está en la imagen de g.

(a) $\mathbf{x}_0 = (0, 0, 0)$, $\mathbf{x}_1 = (1, 0, 1)$, $\mathbf{x}_3 = (2, 3, 0)$; $\mathbf{a} = (4, 3, 1)$.
(b) $\mathbf{x}_0 = (1, -1, 1)$, $\mathbf{x}_1 = (1, 1, 0)$, $\mathbf{x}_2 = (-1, 2, 0)$; $\mathbf{a} = (1, 1, 2)$.
(c) $\mathbf{x}_0 = (0, 0, 1)$, $\mathbf{x}_1 = (1, 2, 1)$, $\mathbf{x}_2 = (2, 1, 2)$; $\mathbf{a} = (3, 0, 4)$.

3. Dibuje las rectas en $\mathcal{R}^2$ 0 $\mathcal{R}^3$ determinadas por

(a) $t(1, 2) + (-1, -1)$.
(b) $t(1, 0, 1) + (1, 1, 0)$.
(c) $(1, 2) \cdot \mathbf{x} = 2$.

4. Dibuje los planos en $\mathcal{R}^3$ determinados por

(a) $u(1, 2, 0) + v(2, 0, 1) + (1, 0, 0)$.
(b) $(-1, -1, 1) \cdot \mathbf{x} = 1$.
(c) $2x + y + z = 0$.

5. Encuentre una ecuación para la recta en $\mathcal{R}^2$ perpendicular a la recta $t\mathbf{x}_1 + \mathbf{x}_0$ que pasa por el punto $\mathbf{y}_2$.

6. Encuentre el coseno del ángulo entre los planos $\mathbf{x}_1 \cdot \mathbf{x} = c_1$ e $\mathbf{y}_1 \cdot \mathbf{y} = c_2$.

7. Demuestre que si $\mathbf{x}_1 \cdot \mathbf{x} = c_1$ e $\mathbf{y}_1 \cdot \mathbf{y} = c_2$ son las ecuaciones normalizadas de dos planos, entonces el coseno del ángulo entre ellos es $\mathbf{x}_1 \cdot \mathbf{y}_1$.

8. Para cada uno de los puntos y planos o rectas indicados más abajo, encuentre la distancia entre el punto y el plano o recta.

(a) $(1, 0, -1)$, $(1, 1, 1) \cdot \mathbf{x} = 1$.
(b) $(1, 0, -1)$, $x + 2y + 3z = 1$.
(c) $(1, 2)$, $(3, 4) \cdot \mathbf{x} = 0$.

9. (a) Verifique que el producto cruz de $\mathbf{x}_1$ y $\mathbf{x}_2$ [véase Ec. (2)], es perpendicular a $\mathbf{x}_2$.
(b) Encuentre una representación de la recta perpendicular al plano formado por todos los puntos $u(1, 2, 1) + v(-1, 0, 1) + (1, 1, 1)$ y que pasa por el origen.

10. (a) Encuentre una representación para el punto $\mathbf{x} = (x, y, z)$ que satisface

$$2x + 3y + z = 1$$

encontrando vectores $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_0$ tales que $\mathbf{x} = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$. [*Sugerencia*: Tome $u = y$, $v = z$.]

(b) Haga lo mismo que en la parte (a) para la ecuación general $\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$ si $\mathbf{p} \neq 0$. [*Sugerencia*: Si $\mathbf{p} \neq 0$, entonces $\mathbf{p}$ tiene una coordenada distinta de cero.]

11. (a) Si $\mathbf{x}$ es un vector en $\mathcal{R}^3$, demuestre que

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x} \cdot \mathbf{e}_2)\mathbf{e}_2 + (\mathbf{x} \cdot \mathbf{e}_3)\mathbf{e}_3.$$

(b) Si el vector $\mathbf{x}$ de la parte (a) es unitario; es decir, un vector $\mathbf{u}$ de largo 1, demuestre que $\mathbf{u} \cdot \mathbf{e}_i = \cos \alpha_i$, donde α_i es el ángulo entre $\mathbf{u}$ y $\mathbf{e}_i$. Las coordenadas $\cos \alpha_i$ se llaman los **cosenos direccionales** de $\mathbf{u}$ con respecto de la base natural. Si $\mathbf{x}$ es un vector distinto de cero, los cosenos direccionales de $\mathbf{x}$ son, por definición, los del vector unitario $\mathbf{x}/|\mathbf{x}|$.
(c) Encuentre los cosenos direccionales de $(1, 2, 1)$.
(d) Demuestre que las partes (a) y (b) se generalizan a $\mathcal{R}^n$.

12. Sean $\mathbf{u}$ y $\mathbf{v}$ puntos de $\mathcal{R}^n$. Demuestre que el punto $\frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$ está en la mitad del segmento de recta que une a $\mathbf{u}$ y $\mathbf{v}$.

13. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores no colineales en $\mathcal{R}^3$. Para encontrar un vector $\mathbf{x}$ perpendicular a $\mathbf{u}$ y a $\mathbf{v}$, resolvemos las ecuaciones $\mathbf{u} \cdot \mathbf{x} = 0$ y $\mathbf{v} \cdot \mathbf{x} = 0$, es decir, resolvemos

$$\begin{aligned} u_1x + u_2y + u_3z &= 0 \\ v_1x + v_2y + v_3z &= 0, \end{aligned} \qquad (*)$$

donde $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ y $\mathbf{x} = (x, y, z)$.
(a) Demuestre que

$$(u_2v_1 - u_1v_2)y = (u_1v_3 - u_3v_1)z$$
$$(u_1v_2 - u_2v_1)x = (u_2v_3 - u_3v_2)z.$$

(b) Demuestre que si $u_1v_2 - u_2v_2 \neq 0$, entonces las ecuaciones $(*)$ tienen una solución

$$(x, y, z) = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1).$$

(c) Indique cómo resolver $(*)$ si $u_1v_2 - u_2v_1 = 0$.

7. Determinantes, definiciones y propiedades geométricas

En esta sección definiremos y estudiaremos las propiedades de cierta función real definida en el conjunto de todas las matrices cuadradas. El valor de esta función para una matriz cuadrada M se llama el **determinante** de M y se escribe $\det M$. Otra manera habitual de denotar el determinante de una matriz, cuando ésta está expresada por sus elementos, es la de remplazar los paréntesis por barras verticales. Así, las notaciones

$$\begin{vmatrix} 1 & 4 & 5 \\ 6 & 7 & -3 \\ -2 & 1 & 0 \end{vmatrix}, \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

significan lo mismo que

$$\det \begin{pmatrix} 1 & 4 & 5 \\ 6 & 7 & -3 \\ -2 & 1 & 0 \end{pmatrix}, \qquad \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Nuestra definición del determinante será inductiva; es decir, definiremos $\det M$ primero para matrices 1-por-1 y luego, para cada n, definiremos el determinante de una matriz n-por-n en términos de los determinantes de ciertas matrices $(n-1)$-por-$(n-1)$. Primero, necesitamos una notación para ciertas submatrices de una matriz dada. Si A es una matriz cualquiera, la matriz obtenida quitando la fila i y la columna j se llama el **menor** ij de A y se denota por A_{ij}. (Recuerde que usamos la letra minúscula a_{ij} para denotar el elemento ij de una matriz A.)

Ejemplo 1. Sean

$$A = \begin{pmatrix} -5 & -6 & 7 \\ 8 & -9 & 0 \\ -3 & 4 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Entonces

$$a_{11} = -5, \qquad A_{11} = \begin{pmatrix} -9 & 0 \\ 4 & 2 \end{pmatrix}$$

$$a_{23} = 0, \qquad A_{23} = \begin{pmatrix} -5 & -6 \\ -3 & 4 \end{pmatrix}$$

$$b_{11} = 1, \qquad B_{11} = (4), \quad b_{12} = 2, \quad B_{12} = (3).$$

Ahora podemos definir la noción de **determinante.**

Para una matriz 1-por-1 $A = (a)$, definimos

$$\det A = a.$$

Para una matriz n-por-n $A = (a_{ij})$, $i, j = 1, \ldots, n$, definimos

$$\textbf{7.1} \qquad \det A = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det A_{1j}$$

$$= a_{11} \det A_{11} - a_{12} \det A_{12} + \ldots - (-1)^n a_{1n} \det A_{1n}.$$

En palabras, esta fórmula dice que $\det A$ es la suma, con signos alternantes de los elementos de la primera fila, cada uno de los cuales está multiplicado por el determinante del menor que le corresponde.

Por esta razón, los números

$$\det A_{11}, \ -\det A_{12}, \ \ldots, \ (-1)^{n+1} \det A_{1n}$$

se llaman los **cofactores** de los elementos correspondientes de la primera fila de A. En general, el **cofactor** del elemento a_{ij} de A es, por definición, $(-1)^{i+j} \det A_{ij}$. Así, en el Ejemplo 1, el elemento $a_{21} = 8$ de la matriz A tiene el cofactor

$$(-1)^{2+1} \det \begin{pmatrix} -6 & 7 \\ 4 & 2 \end{pmatrix} = 40.$$

El factor $(-1)^{i+j}$ asocia signos más y menos a $\det A_{ij}$ de acuerdo con el esquema siguiente:

$$\begin{pmatrix} + & - & + & - & \cdot & \cdot & \cdot \\ - & + & - & + & \cdot & \cdot & \cdot \\ + & - & + & - & \cdot & \cdot & \cdot \\ - & + & - & + & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \end{pmatrix}$$

Ejemplo 2.

(a) $$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1(4) - 2(3) = 4 - 6 = -2.$$

(b) $$\det \begin{pmatrix} -5 & -6 & 7 \\ 8 & -9 & 0 \\ -3 & 4 & 2 \end{pmatrix} = -5 \det \begin{pmatrix} -9 & 0 \\ 4 & 2 \end{pmatrix} - (-6) \det \begin{pmatrix} 8 & 0 \\ -3 & 2 \end{pmatrix} + 7 \det \begin{pmatrix} 8 & -9 \\ -3 & 4 \end{pmatrix}$$

$$= -5(-18 - 0) + 6(16 - 0) + 7(32 - 27)$$

$$= 90 + 96 + 35 = 221.$$

(c) $$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Es conveniente recordar el resultado del último ejemplo como una regla para calcular. *El determinante de una matriz 2-por-2 es el producto de los elementos de la diagonal principal menos el producto de los otros elementos.* Por tanto, el determinante de una matriz 2-por-2 puede generalmente ser calculado mentalmente y, el de una matriz 3-por-3, en una o dos líneas. En principio, cualquier determinante puede ser calculado a partir de la definición, pero esto involucra números grandes de operaciones aritméticas si la dimensión es un poco grande. Algunos de los teoremas que demostraremos justificarán otros métodos de cálculo con menos operaciones aritméticas que las que se requieren en el caso $n > 3$ si se parte directamente de la definición.

Los determinantes fueron inventados (en la mitad del siglo xviii) como un medio para expresar la solución de sistemas de ecuaciones lineales. Para ver cómo funciona esto para dos ecuaciones con dos incógnitas, consideremos el sistema general

$$a_{11}x_1 + a_{12}x_2 = r_1$$

$$a_{21}x_1 + a_{22}x_2 = r_2.$$

La variable x_2 puede ser eliminada multiplicando la primera ecuación por a_{22} y la segunda por a_{12} y tomando la diferencia. El resultado es la ecuación

$$(a_{22}a_{11} - a_{12}a_{21})x_1 = (a_{22}r_1 - a_{12}r_2).$$

que puede ser escrita en la forma

$$x_1 \det A = \det B^{(1)}$$

donde $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ es la matriz de coeficientes y $B^{(1)} = \begin{pmatrix} r_1 & a_{12} \\ r_2 & a_{22} \end{pmatrix}$ es la

matriz obtenida remplazando la primera columna de A por $\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$. El lector podrá obtener fácilmente la ecuación $x_2 \det A = \det B^{(2)}$, donde $B^{(2)} = \begin{pmatrix} a_{11} & r_1 \\ a_{21} & r_2 \end{pmatrix}$ es el resultado de remplazar la segunda columna de A por $\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$. En el Teorema 8.4 daremos un resultado general para el caso de n ecuaciones lineales con n incógnitas, para todo n.

Puesto que la definición de determinantes dada en 7.1 es inductiva, la mayor parte de las demostraciones tienen el mismo carácter. Es decir, para demostrar un teorema sobre determinantes, lo verificamos para matrices 1-por-1 (o, en algunos casos, para matrices 2-por-2) y demostramos también que si es válido para matrices $(n-1)$-por-$(n-1)$, también es válido para matrices n-por-n. En las demostraciones daremos solo el argumento para pasar de $n-1$ a n. El lector deberá verificar las proposiciones directamente para matrices 1-por-1 y para 2-por-2; la verificación es siempre trivial. Denotaremos matrices n-por-n por las letras A, B, C. Escribiremos $\mathbf{a}_j$ para la columna j de A. Si A tiene n filas, entonces $\mathbf{a}_j$ es un vector en $\mathcal{R}^n$.

7.2 Teorema

Si B se obtiene de A por multiplicación de alguna columna por un número r, entonces $\det B = r \det A$.

Demostración. Supongamos que $\mathbf{b}_j = r\mathbf{a}_j$, mientras que $\mathbf{b}_k = \mathbf{a}_k$ para $k \neq j$. Entonces, en particular, $b_{1j} = ra_{1j}$. Si $k \neq j$, B_{1k} se obtiene a partir de A_{1k} multiplicando una columna por r y como B_{1k} y A_{1k} son matrices $(n-1)$-por-$(n-1)$, tenemos que $\det B_{1k} = r \det A_{1k}$. Por otra parte, $B_{1j} = A_{1j}$ y $b_{1k} = a_{1k}$ si $k \neq j$. Por tanto, para todo k, $b_{1k} \det B_{1k} = ra_{1k} \det A_{1k}$. Por tanto,

$$\det B = \sum_{k=1}^{n} (-1)^{k+1} b_{1k} \det B_{1k}$$

$$= \sum_{k=1}^{n} (-1)^{k+1} r a_{1k} \det A_{1k} = r \det A.$$

7.3 Corolario

Si una matriz tiene una columna cero, su determinante es cero.

Demostración. Si $\mathbf{a}_j = 0$, entonces $\mathbf{a}_j = 0 \cdot \mathbf{a}_j$. Entonces, por el Teorema 7.2, $\det A = 0 \cdot \det A = 0$.

Ejemplo 3. Sean

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 4 \\ 0 & 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 6 & 3 \\ -1 & 6 & 4 \\ 0 & 3 & 2 \end{pmatrix}.$$

B se obtiene de A multiplicando la segunda columna por 3.

$$\det A = (1)(4-4) - 2(-2-0) + 3(-1+0)$$
$$= 0 + 4 - 3 = 1$$
$$\det B = (1)(12-12) - 6(-2-0) + 3(-3+0)$$
$$= 0 + 12 - 9 = 3 = 3 \det A.$$

7.4 Teorema

Sean A, B y C iguales, excepto para la columna j, y supongamos que la columna j de C es la suma de las columnas j de A y B. Entonces $\det C = \det A + \det B$.

Demostración. Tenemos que $c_{1j} = a_{1j} + b_{1j}$ y $C_{1j} = A_{1j} = B_{1j}$. Para $k \neq J$, $c_{1k} = a_{1k} = b_{1k}$ y C_{1k} es igual a A_{1k} y B_{1k}, excepto para una columna que es la suma de las columnas correspondientes de A_{1k} y B_{1k}. Por tanto, si $k \neq j$, $\det C_{1k} = \det A_{1k} + \det B_{1k}$ por hipótesis de inducción. Si $k = j$ tenemos

$$c_{1k} \det C_{1k} = c_{1k} \det A_{1k} + c_{1k} \det B_{1k}$$
$$= a_{1k} \det A_{1k} + b_{1k} \det B_{1k},$$

mientras que

$$c_{1j} \det C_{1j} = a_{1j} \det C_{1j} + b_{1j} \det C_{1j}$$
$$= a_{1j} \det A_{1j} + b_{1j} \det B_{1j}.$$

Por tanto,

$$\det C = \sum_{k=1}^{n} (-1)^{k+1} c_{1k} \det C_{1k}$$

$$= \sum_{k=1}^{n} (-1)^{k+1} a_{1k} \det A_{1k} + \sum_{k=1}^{n} (-1)^{k-1} b_{1k} \det B_{1k}$$

$$= \det A + \det B.$$

Ejemplo 4. Sean

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 4 \\ 0 & 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 3 & 3 \\ -1 & 1 & 4 \\ 0 & -2 & 2 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 5 & 3 \\ -1 & 3 & 4 \\ 0 & -1 & 2 \end{pmatrix}.$$

A, B y C son iguales, excepto para las segundas columnas, y la segunda columna de C es la suma de las segundas columnas de A y de B.

$$\begin{aligned}\det A &= (1)(4-4) - 2(-2-0) + 3(-1-0)\\ &= 0 + 4 - 3 = 1\\ \det B &= (1)(2+8) - 3(-2-0) + 3(2-0)\\ &= 10 + 6 + 6 = 22\\ \det C &= (1)(6+4) - 5(-2-0) + 3(1-0)\\ &= 10 + 10 + 3 = 23 = \det A + \det B.\end{aligned}$$

Los Teoremas 7.2 y 7.4 tienen una interpretación importante si se considera al determinante de una matriz como una función de las columnas de la matriz. De esta manera, si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ son las columnas de una matriz n-por-n, podemos escribir $\det A = \det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Mantengamos fijos los $(n-1)$ vectores $\mathbf{a}_2, \ldots, \mathbf{a}_n$ y consideremos la función de $\mathbf{x}$, $f(\mathbf{x}) = \det(\mathbf{x}, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Entonces, los Teoremas 7.2 y 7.4 afirman que $f(r\mathbf{x}) = rf(\mathbf{x})$ y $f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2)$. Lo mismo sería válido si el vector «variable» estuviese en el lugar j en lugar del primero. En otras palabras, tenemos

7.5 Teorema

Si todos menos uno de entre $\mathbf{a}_1, \ldots, \mathbf{a}_n$ se mantiene fijo, entonces $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ es una función lineal del vector restante.

Para ver cómo funciona esto para la primera columna de una matriz, sea

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Para expresar $\det(\mathbf{x}, \mathbf{b}, \mathbf{c})$ como una función lineal en $\mathbf{x}$, necesitamos simplemente calcular

$$\begin{aligned}\det(\mathbf{x}, \mathbf{b}, \mathbf{c}) &= \det \begin{pmatrix} x_1 & b_1 & c_1 \\ x_2 & b_2 & c_2 \\ x_3 & b_3 & c_3 \end{pmatrix}\\ &= x_1 \det \begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} - b_1 \det \begin{pmatrix} x_2 & c_2 \\ x_3 & c_3 \end{pmatrix} + c_1 \det \begin{pmatrix} x_2 & b_2 \\ x_3 & b_3 \end{pmatrix}\\ &= x_1(b_2c_3 - b_3c_2) - b_1(x_2c_3 - x_3c_2) + c_1(x_2b_3 - x_3b_2)\\ &= x_1(b_2c_3 - b_3c_2) - x_2(b_1c_3 - b_3c_1) + x_3(b_1c_2 - c_1b_2),\end{aligned}$$

lo cual es una función lineal de $\mathbf{x}$.

Tan importante como la condición de linealidad para los determinantes es el hecho de que si se intercambian dos columnas (o filas) en una matriz, el determinante cambia de signo. Empezamos demostrando el resultado para columnas *adyacentes*.

7.6 Lema

Si B se obtiene a partir de A intercambiando dos columnas adyacentes, entonces $\det B = -\det A$.

Demostración. Supongamos que A y B son iguales, excepto que $\mathbf{a}_j = \mathbf{b}_{j+1}$ y $\mathbf{a}_{j+1} = \mathbf{b}_j$. Si $k \neq j$ y $j+1$, tenemos que $b_{1k} = a_{1k}$ y $\det B_{1k} = -\det A_{1k}$ por la hipótesis de inducción, de modo que $(-1)^{k+1} b_{1k} \det B_{1k} = -(-1)^{k+1} a_{1k} \det A_{1k}$. Por otra parte, $b_{1j} = a_{1,j+1}$, $B_{1j} = A_{1,j+1}$, de modo que $(-1)^{j+1} b_{1j} \det B_{1j} = (-1)^{j+1} a_{1,j+1} \det A_{1,j+1} = -(-1)^{j+2} a_{1,j+1} \det A_{1,j+1}$. De la misma manera, $(-1)^{j+2} b_{1,j+1} = -(-1)^{j+1} a_{1j} \det A_{1j}$. Por tanto, cada término de la expresión para $\det B$, dada en 7.1, es igual al negativo de un término en la expresión para $\det A$. Por tanto, $\det B = -\det A$.

Ejemplo 5. Sean

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & 5 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & -4 \\ 3 & -2 & 5 \end{pmatrix}.$$

Entonces

$$\begin{aligned}\det A &= (1)(8-5) - (3)(-4-3) + (-2)(10-(-12))\\ &= 3 + 21 - 44 = -20\\ \det B &= (1)(5-8) - (-2)(10-(-12)) + (3)(-4-3)\\ &= -3 + 44 - 21 = 20\\ &= -\det A\end{aligned}$$

7.7 Teorema

Si B se obtiene a partir de A intercambiando dos columnas, entonces $\det B = -\det A$.

Demostración. Supongamos que hay k columnas entre las columnas en cuestión (de modo que $k = 0$ si ellas son adyacentes). La primera columna puede hacerse adyacente a la segunda haciendo k intercambios entre columnas adyacentes; entonces se pueden intercambiar las dos columnas y con k intercambios de columnas adyacentes, la segunda columna puede ponerse en el lugar en que estaba la primera. El número total de intercambios de columnas adyacentes es $2k+1$ y, por el Lema 7.6, cada intercambio cambia el signo del determinante. Como $2k+1$ es un número impar, $\det B = -\det A$.

7.8 Teorema

Si dos columnas de A son iguales, entonces $\det A = 0$.

Demostración. Intercambiando las dos columnas iguales de A, se obtiene nuevamente A. Entonces $\det A = -\det A$ y $\det A = 0$.

La multiplicación por una matriz n-por-n define una función lineal de $\mathcal{R}^n$ en $\mathcal{R}^n$ y es natural preguntarse si el determinante de la matriz está relacionado con alguna propiedad de la función lineal asociada. Resulta que el determinante describe cómo esta función afecta volúmenes en $\mathcal{R}^n$. En $\mathcal{R}^2$, por supuesto, «volumen» significa área.

Ejemplo 6. La multiplicación por $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ define una función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^2$ que multiplica por 3 las longitudes en la dirección x y por 2 las longitudes en la dirección y. Todas las áreas son magnificadas por un factor de 6, como lo ilustra la Fig. 24(a), que muestra un cuadrado unitario S y su imagen $f(S)$. Nótese que $\det \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} = 6$. Otro ejemplo es la función g con matriz $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, que tiene determinante 1. Su efecto se ilustra en la Fig. 24(b). El cuadrado unitario se transforma en un paralelogramo con la misma base y altura, de modo que el área no cambia. La compuesta $g \circ f$ multiplica las áreas por 6 [ya que $f(S)$ tiene 6 veces el área de S y $g(f(S))$ tiene la misma área que $f(S)$]. La matriz de $g \circ f$ está dada por la matriz producto

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix},$$

y vemos que su determinante es 6.

Estos ejemplos sugieren que una función lineal de $\mathcal{R}^2$ en $\mathcal{R}^2$ multiplica las áreas por un factor igual al determinante de la matriz asociada. Esto no es totalmente cierto. Por ejemplo, es claro que la reflexión en el eje y (que envía el punto (x, y) al punto $(-x, y)$) no altera las áreas mientras que su matriz $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ tiene determinante -1. Para obtener una afirmación correcta, debemos usar el valor absoluto del determinante. En efecto, la siguiente afirmación general es verdadera.

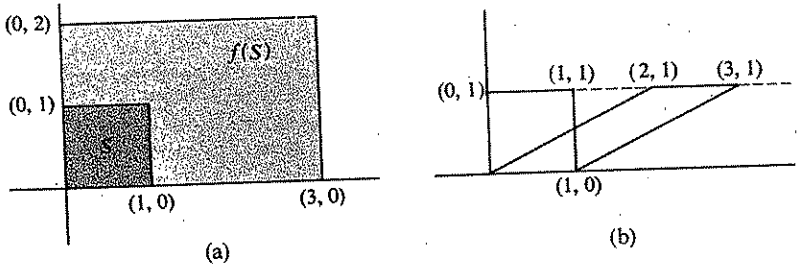


Figura 24

7.9 Teorema

Una función lineal de $\mathcal{R}^n$ en $\mathcal{R}^n$ con matriz A multiplica los volúmenes por el factor $|\det A|$.

Este teorema no puede ser demostrado, en esta forma general, sin una definición rigurosa de volumen. Por tanto, la demostración será dada en el Teorema 3.2 del Cap. 4, donde se obtiene como un caso especial del teorema demostrado en la Sec. 9 del Apéndice.

Si una función altera los volúmenes en un factor de u y otra en un factor de v, entonces la función compuesta altera volúmenes en un factor de uv. Por tanto, el Teorema 7.9 implica que para matrices n-por-n, A, B, se tiene que $|\det (AB)| = |\det A| \cdot |\det B|$. En efecto, el siguiente teorema es válido.

7.10 Regla del producto

Si A y B son dos matrices cuadradas del mismo tamaño, entonces $\det (AB) = (\det A)(\det B)$.

Demostración. Sea

$$L(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det A \det (\mathbf{x}_1, \ldots, \mathbf{x}_n) - \det (A\mathbf{x}_1, \ldots, A\mathbf{x}_n),$$

donde $\mathbf{x}_1, \ldots, \mathbf{x}_n$ son vectores en $\mathcal{R}^n$. Es claro que L es lineal como función de cada $\mathbf{x}_j$. Además, $L(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = 0$ para cualquier conjunto $\{\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}\}$ de vectores de la base natural de $\mathcal{R}^n$.

La razón para esto es que si hay algún vector repetido, entonces $\det (\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n})$ y $\det (A\mathbf{e}_{i_1}, \ldots, A\mathbf{e}_{i_n})$ son cero, por el Teorema 7.8. Si esto no es el caso, entonces $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}$ son $\mathbf{e}_1, \ldots, \mathbf{e}_n$ en algún orden y por el Teorema 7.7

$$\begin{aligned} \det A \det (\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) &= \pm \det A \det (\mathbf{e}_1, \ldots, \mathbf{e}_n) = \pm \det A \det I \\ &= \pm \det A = \pm \det (A\mathbf{e}_1, \ldots, A\mathbf{e}_n) \\ &= \det (A\mathbf{e}_{i_1}, \ldots, A\mathbf{e}_{i_n}). \end{aligned}$$

Pero entonces $L(\mathbf{b}_1, \ldots, \mathbf{b}_n) = 0$, donde

$$\mathbf{b}_j = \begin{pmatrix} b_{1j} \\ \cdot \\ \cdot \\ \cdot \\ b_{nj} \end{pmatrix}$$

es la columna j de B. En efecto, usando la linealidad de L,

$$L(\mathbf{b}_1, \ldots, \mathbf{b}_n) = L\left(\sum_{i=1}^{n} b_{i1}\mathbf{e}_i, \ldots, \sum_{i=1}^{n} b_{in}\mathbf{e}_i\right)$$

$$= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1 1} \cdots b_{i_n n} L(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = 0.$$

Por tanto,

$$\det A \det B - \det AB = \det A \det (\mathbf{b}_1, \ldots, \mathbf{b}_n) - \det (A\mathbf{b}_1, \ldots, A\mathbf{b}_n) = L(\mathbf{b}_1, \ldots, \mathbf{b}_n) = 0.$$

Nótese que la demostración solo usa los Teoremas 7.5, 7.7 y 7.8 y que no usa el Teorema 7.9. (Esto es importante, ya que la regla producto se usa en la Sec. 7 del Apéndice, en una demostración de la cual depende la demostración del Teorema 7.9.)

La unidad natural de área en $\mathcal{R}^2$ está dada por el cuadrado unitario con lados (1, 0) y (0, 1) y la unidad natural de volumen en $\mathcal{R}^3$ está dada por el cubo unitario de lados (1, 0, 0), (0, 1, 0), (0, 0, 1). En general, tomamos como unidad de volumen en $\mathcal{R}^n$ al volumen del cubo unitario cuyos ejes son los vectores de la base natural: $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Además, podemos tomar n vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ en $\mathcal{R}^n$ y formar todas las combinaciones lineales $t_1\mathbf{x}_1 + \ldots + t_n\mathbf{x}_n$, donde cada uno de los números $t_1, \ldots, t_n$ satisface la condición $0 \leq t_i \leq 1$. El conjunto de puntos resultante se llama el **paralelepípedo** determinado por sus lados $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Si elegimos solo dos vectores $\mathbf{x}_1$, $\mathbf{x}_2$, entonces hablamos del **paralelogramo** determinado por $\mathbf{x}_1$ y $\mathbf{x}_2$. La Fig. 25 muestra un paralelepípedo.

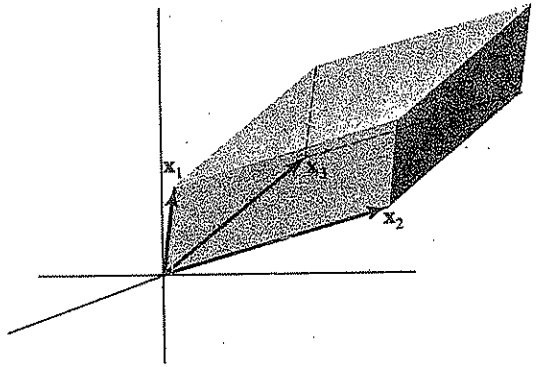


Figura 25

7.11 Teorema

Sean $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, n vectores en $\mathcal{R}^n$. Entonces el volumen del paralelepípedo de lados $\mathbf{a}_1, \ldots, \mathbf{a}_n$ es $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|$.

Demostración. La función lineal cuya matriz tiene columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$ transforma $\mathbf{e}_i$ en $\mathbf{a}_j$ (por el Teorema 4.2). Por tanto, f transforma el cubo unitario en el paralelepípedo de lados a_j. Puesto que f multiplica volúmenes por el factor $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|$ y el cubo tiene volumen unidad, el volumen del paralelepípedo es $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|$.

Ejemplo 7. Sea $\mathbf{a}_1 = \begin{pmatrix} r_1 \cos\theta_1 \\ r_1 \operatorname{sen}\theta_1 \end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix} r_2 \cos\theta_2 \\ r_2 \operatorname{sen}\theta_2 \end{pmatrix}$. Estos vectores tienen largos r_1, r_2 y forman ángulos θ_1, θ_2 con el eje x, como es indicado en la Fig. 26. Entonces $\det(\mathbf{a}_1, \mathbf{a}_2)$ es

$$\det \begin{pmatrix} r_1 \cos\theta_1 & r_2 \cos\theta_2 \\ r_1 \operatorname{sen}\theta_1 & r_2 \operatorname{sen}\theta_2 \end{pmatrix} = r_1 r_2 (\cos\theta_1 \operatorname{sen}\theta_2 - \operatorname{sen}\theta_1 \cos\theta_2)$$

$$= r_1 r_2 \operatorname{sen}(\theta_2 - \theta_1)$$

Este número puede ser interpretado como el producto de la base r_1 por la altura perpendicular $r_2 \operatorname{sen}(\theta_2 - \theta_1)$.

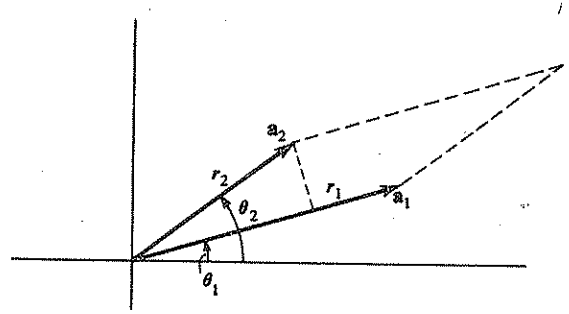


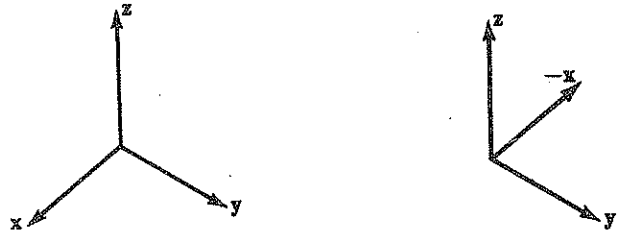
Figura 26

Hemos visto que el valor absoluto de $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ puede ser interpretado como un volumen. El signo de este determinante también tiene una interpretación geométrica. Decimos que un conjunto ordenado $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ de vectores en $\mathcal{R}^n$ tiene **orientación positiva** (o está **orientado positivamente**) si $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) > 0$ y que tiene **orientación negativa** si $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) < 0$. Si el determinante es cero, la orientación no está definida.

El Ejemplo 7 muestra que, en $\mathcal{R}^2$, el signo de $\det(\mathbf{a}_1, \mathbf{a}_2)$ es el mismo que el signo de $\operatorname{sen}\theta$, donde $\theta = \theta_2 - \theta_1$ es el ángulo de $\mathbf{a}_1$ a $\mathbf{a}_2$. Por tanto, la orientación de $(\mathbf{a}_1, \mathbf{a}_2)$ es positiva si es posible llevar a $\mathbf{a}_1$ a la misma dirección que $\mathbf{a}_2$ mediante una rotación de menos de 180° en el sentido contrario al de los punteros del reloj. La orientación es negativa si se necesita una rotación en el sentido de los punteros del reloj. No está definida si $\mathbf{a}_1$ y $\mathbf{a}_2$ están en una recta

común. Es decir, en $\mathcal{R}^2$, la orientación corresponde a la dirección de rotación. (Nótese que la orientación de $(\mathbf{a}_1, \mathbf{a}_2)$ es opuesta a la de $(\mathbf{a}_2, \mathbf{a}_1)$.)

La propiedad 7.7 de los determinantes significa que la orientación de un conjunto de vectores cambia cuando se intercambian dos vectores del conjunto. La interpretación de la orientación en $\mathcal{R}^3$ es menos obvia. Los conjuntos $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ y $(-\mathbf{x}, \mathbf{y}, \mathbf{z})$, mostrados en las Figs. 27(a) y 27(b), tienen orientaciones opuestas, ya que, por el Teorema 7.2, $\det(-\mathbf{x}, \mathbf{y}, \mathbf{z}) = -\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$. El conjunto ordenado $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ se llama un sistema derecho porque si se hace que los dedos pulgar e índice de la mano derecha indiquen en las direcciones de $\mathbf{x}$ e $\mathbf{y}$, respectivamente, entonces el dedo medio indicará en la dirección de $\mathbf{z}$. De manera análoga, $(-\mathbf{x}, \mathbf{y}, \mathbf{z})$ forman un sistema izquierdo, ya que se puede hacer lo mismo con la mano izquierda. En este libro hemos preferido dibujar los vectores en el espacio de tal manera que los vectores $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ formen un sistema derecho. Como $\det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \det I = 1$, esto implica que nuestros sistemas derechos tienen orientación positiva y que los sistemas izquierdos tienen orientación negativa.



(a) sistema derecho (b) sistema izquierdo

Figura 27

Sean $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$ vectores en $\mathcal{R}^3$. El vector con coordenadas

$$\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, \quad \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix}, \quad \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$$

se llama el **producto cruz** de $\mathbf{u}$ y $\mathbf{v}$ y se denota por $\mathbf{u} \times \mathbf{v}$. Nótese que al intercambiar $\mathbf{u}$ y $\mathbf{v}$, el signo de cada una de las componentes cambia de modo que $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$. El significado del producto cruz está en la ecuación

7.12 $$\mathbf{x} \cdot (\mathbf{u} \times \mathbf{v}) = \det(\mathbf{x}, \mathbf{u}, \mathbf{v}),$$

que es válida para todo $\mathbf{x}$ en $\mathcal{R}^3$. Para demostrar esto, sea $\mathbf{x} = (x_1, x_2, x_3)$. Entonces

$$\det(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \det \begin{pmatrix} x_1 & u_1 & v_1 \\ x_2 & u_2 & v_2 \\ x_3 & u_3 & v_3 \end{pmatrix}$$

$$\begin{aligned}
&= x_1 \begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix} - u_1 \begin{vmatrix} x_2 & v_2 \\ x_3 & v_3 \end{vmatrix} + v_1 \begin{vmatrix} x_2 & u_2 \\ x_3 & u_3 \end{vmatrix} \\
&= x_1(u_2v_3 - u_3v_2) - u_1(x_2v_3 - x_3v_2) + v_1(x_2u_3 - x_3u_2) \\
&= x_1(u_2v_3 - u_3v_2) + x_2(u_3v_1 - u_1v_3) + x_3(u_1v_2 - u_2v_1) \\
&= x_1 \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} + x_2 \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} + x_3 \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \\
&= \mathbf{x} \cdot (\mathbf{u} \times \mathbf{v}).
\end{aligned}$$

Una manera conveniente de recordar la fórmula para el producto cruz es pensar en el «determinante» formal

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{e}_3$$

que expresa $\mathbf{u} \times \mathbf{v}$ en términos de la base estándar $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Si $\mathbf{u}$ o $\mathbf{v}$ es cero, o si uno es un múltiplo numérico del otro, claramente $\mathbf{u} \times \mathbf{v} = 0$. En caso contrario, $\mathbf{u}$ y $\mathbf{v}$ determinan un plano. Si sustituimos $\mathbf{u}$ y $\mathbf{v}$ por $\mathbf{x}$ en 7.12, tenemos

$$\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = \det(\mathbf{u}, \mathbf{u}, \mathbf{v}) = 0, \qquad \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = \det(\mathbf{v}, \mathbf{u}, \mathbf{v}) = 0$$

por el Teorema 7.8. Por tanto, $\mathbf{u} \times \mathbf{v}$ es perpendicular a $\mathbf{u}$ y a $\mathbf{v}$. Sea $\mathbf{x}$ un vector unitario en la dirección de $\mathbf{u} \times \mathbf{v}$. Entonces $\mathbf{x} \cdot (\mathbf{u} \times \mathbf{v}) = |\mathbf{u} \times \mathbf{v}|$. Por otra parte, $\mathbf{x} \cdot (\mathbf{u} \times \mathbf{v}) = \det(\mathbf{x}, \mathbf{u}, \mathbf{v})$, y tiene valor absoluto igual al volumen del paralelepípedo de lados $\mathbf{x}$, $\mathbf{u}$, $\mathbf{v}$. Como $\mathbf{x}$ tiene largo unidad y es perpendicular al plano de $\mathbf{u}$ y $\mathbf{v}$, este volumen es igual al área del paralelogramo de lados $\mathbf{u}$, $\mathbf{v}$. Finalmente,

$$\det(\mathbf{u} \times \mathbf{v}, \mathbf{u}, \mathbf{v}) = (\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v}) = |\mathbf{u} \times \mathbf{v}|^2 \geq 0.$$

Por tanto, salvo cuando $\mathbf{u} \times \mathbf{v} = 0$, el triple $(\mathbf{u} \times \mathbf{v}, \mathbf{u}, \mathbf{v})$ tiene orientación positiva y forma un sistema derecho.

Podemos resumir lo que hemos demostrado como sigue:

7.13 Teorema

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores no colineales en $\mathcal{R}^3$, entonces su producto cruz $\mathbf{u} \times \mathbf{v}$ es un vector perpendicular a $\mathbf{u}$ y $\mathbf{v}$ y de largo igual al área del paralelogramo de lados $\mathbf{u}$ y $\mathbf{v}$. El triple $(\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v})$ es un sistema derecho.

La Fig. 28 muestra la relación entre $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} \times \mathbf{v}$.

Ejemplo 8. Encontremos el área del paralelogramo de lados $\mathbf{u} = (1, 2, 3)$ y $\mathbf{v} = (3, 2, 1)$. Tenemos

$$\mathbf{u} \times \mathbf{v} = ((2)(1) - (3)(2), -(1)(1) + (3)(3), (1)(2) - (2)(3))$$
$$= (-4, 8, -4).$$
$$\text{Area} = |\mathbf{u} \times \mathbf{v}| = (16 + 64 + 16)^{1/2}$$
$$= \sqrt{96} = 4\sqrt{6}.$$

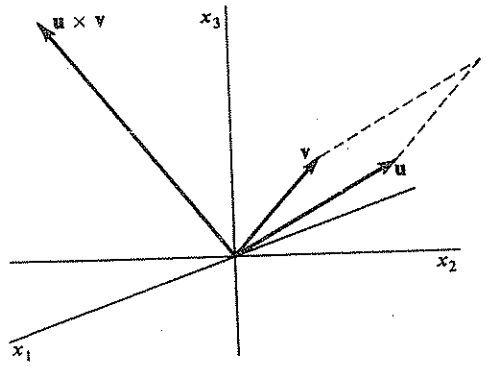


Figura 28

A veces es conveniente combinar las ideas de volumen y orientación en un solo concepto. Definimos el volumen orientado determinado por una n-pla de vectores como el volumen ordinario si la orientación de la n-pla es positiva y su negativo si la orientación es negativa. Entonces, el volumen orientado de $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ es det $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. La relación entre volumen ordinario y volumen orientado es muy parecida a la relación entre la distancia en una recta y la distancia dirigida en la recta. De hecho, el volumen orientado puede ser considerado como una generalización de la noción de distancia dirigida. En el Cap. 7, Sec. 7, haremos uso de esta idea.

EJERCICIOS

1. Encuentre AB, BA y los determinantes de A, B, AB y BA si

(a) $A = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 2 & -3 \end{pmatrix}$. [*Resp.* det $AB = -14$.]

(b) $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 0 & 1 \\ 2 & -1 & -3 \\ 0 & 3 & 5 \end{pmatrix}$.

2. Encuentre los coeficientes necesarios para expresar cada una de las funciones siguientes como una función lineal de los x.

(a) $\begin{vmatrix} x_1 & -1 & 6 \\ x_2 & 4 & -3 \\ x_3 & 2 & 5 \end{vmatrix}$. (b) $\begin{vmatrix} 1 & x_1 & -3 & 2 \\ 0 & x_2 & 0 & 0 \\ 3 & x_3 & -1 & 5 \\ -2 & x_4 & 0 & 1 \end{vmatrix}$.

3. Demuestre que si D es la matriz diagonal **diag** $(r_1, \ldots, r_n)$, entonces det $D = r_1 r_2 \ldots r_n$.

4. ¿Cuál es la relación entre
(a) det A y det $(-A)$?
(b) det $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, \mathbf{a}_n)$ y det $(\mathbf{a}_n, \mathbf{a}_{n-1}, \ldots, \mathbf{a}_2, \mathbf{a}_1)$?

5. Verifique que la regla producto para los pares de matrices en el Ejercicio 1(a) y 1(b).

6. Aplique la regla producto para demostrar que si A es invertible, entonces det $A \neq 0$ y det $(A^{-1}) = (\det A)^{-1}$.

7. Sea A una matriz m-por-n y B una matriz n-por-n. Considere la matriz $(m + n)$-por-$(m + n)$ $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ que contiene a A en su rincón superior derecho, B en su rincón inferior izquierdo y ceros en todas las otras partes. Demuestre que su determinante es (det A) (det B). [*Sugerencia*: Primero considere el caso en que A o B es la matriz identidad, y derive el caso general usando la regla del producto.]

8. Geométricamente es claro que una rotación de $\mathcal{R}^2$ preserva orientaciones y que una reflexión las cambia y que ambas no cambian las áreas. Verifique esto calculando el determinante de sus matrices. (Véanse los Ejercicios 4 y 6 de la Sec. 4.)

9. Usando la interpretación del determinante como un volumen, demuestre que $|\det(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)| \leq |\mathbf{x}_1| \cdot |\mathbf{x}_2| \cdot |\mathbf{x}_3|$ para tres vectores cualesquiera de $\mathcal{R}^3$, y demuestre que hay igualdad si, y solo si, los vectores son ortogonales entre sí.

10. Encuentre el volumen y el área de cada uno de los casos del paralelepípedo con lados $(1, 1, 0)$, $(0, 1, 2)$, $(-3, 5, -1)$.
[*Resp.* volumen = 17, áreas = $\sqrt{166}$, $\sqrt{66}$, 3.]

11. Si $\mathbf{u} = (2, 1, 3)$, $\mathbf{v} = (0, 2, -1)$ y $\mathbf{w} = (1, 1, 1)$, calcule $\mathbf{u} \times \mathbf{v}$, det $(\mathbf{u}, \mathbf{v}, \mathbf{w})$, $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$, $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$ y $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$.

12. Demuestre que $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$, que $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$ y que $(\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (a\mathbf{v}) = a(\mathbf{u} \times \mathbf{v})$, donde a es un número real.

13. Encuentre una representación de la recta que pasa por $(1, 1, 1)$ y es perpendicular a $(2, 1, 3)$ y $(0, 2, -1)$.

14. Sea P el paralelogramo determinado por dos vectores en $\mathcal{R}^3$. Sean P_x, P_y y P_z las proporciones de P en los planos y, z; z, x y x, y respectivamente. Si $A(P)$ es el área de P demuestre que $A^2(P) = A^2(P_x) + A^2(P_y) + A^2(P_z)$.

15. (a) Verifique la relación $|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2 |\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ calculando las coordenadas de los dos términos.
(b) Use el resultado de la parte (a) para demostrar que $|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}| |\mathbf{v}|$ sen θ, donde θ es el ángulo entre $\mathbf{u}$ y $\mathbf{v}$ tal que $0 \leq \theta \leq \pi$.
(c) Demuestre que $|\mathbf{u}| |\mathbf{v}|$ sen θ es el área del paralelogramo de lados $\mathbf{u}$ y $\mathbf{v}$.

16. Los números complejos pueden ser extendidos para formar el álgebra de los cuaterniones $\mathcal{H}$, que es un espacio vectorial de cuatro dimensiones con la base natural $\{1, i, j, k\}$. Así, un cuaternión típico se escribe como $q = a_1 + a_2 i + a_3 j + a_4 k$, donde las a son números reales. Se define un producto $\mathcal{H}$ requiriendo que $i^2 = j^2 = k^2 = -1$ y que $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$. El producto de dos cuaterniones se obtiene multiplicando de la forma habitual y luego usando las reglas citadas para los vectores básicos. Se puede considerar a $\mathcal{R}^3$ como el subespacio $\mathcal{G}$ de $\mathcal{H}$ que consiste de todos los cuaterniones con «parte real» cero y que, tiene como base natural a $\{i, j, k\}$.

(a) Demuestre que el producto cuaterniónico de dos elementos de $\mathcal{G}$ no está necesariamente en $\mathcal{G}$.

(b) Defina un producto en $\mathcal{G}$ de la manera siguiente: Primero se ejecuta el producto de los cuaterniones y luego se remplaza la parte real por cero. Demuestre que esta operación coincide con el producto cruz de $\mathcal{R}^3$.

17. Demuestre la identidad $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ para vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ en $\mathcal{R}^3$. [*Indicación*: Elija un conjunto ortonormal $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ para $\mathcal{R}^3$, de modo que tenga orientación positiva y

$$\mathbf{a} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_3\mathbf{u}_3$$

$$\mathbf{b} = b_1\mathbf{u}_1 + b_2\mathbf{u}_2$$

$$\mathbf{c} = c_1\mathbf{u}_1.]$$

8. Desarrollo de determinantes, regla de Cramer

En la sección anterior definimos el valor de un determinante n-por-n, $A = (a_{ij})$ por

$$\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det A_{1j},$$

donde, en general, A_{ij} es el menor $(n-1)$-por-$(n-1)$ que corresponde a a_{ij}. En esta sección demostraremos algunas fórmulas generales del mismo tipo. Las fórmulas válidas para cualquier matriz n-por-n, A son

8.1
$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

8.1F
$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

La Ec. 8.1 es válida para cualquier valor de j entre 1 y n y la Ec. 8.1F para todo valor de i entre 1 y n. (Nótese que la Ec. 8.1F con $i = 1$ es simplemente la definición original 8.1.) Para cualquier valor de j los elementos a_{ij} de la matriz que aparece en el lado derecho de 8.1 están en la columna j de A. Por esta razón, el término de la derecha de 8.1 se llama el **desarrollo de** det A **por su columna** j. Análogamente, el término de la derecha de 8.1F se llama el **desarrollo de** det A **por la fila** i. Pospondremos la demostración de estas fórmulas, mostrando primero algunas de sus consecuencias.

Ejemplo 1. Sea

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 0 \end{pmatrix}.$$

La expansión de det A por la segunda fila es

$$\begin{aligned} &-5 \det \begin{pmatrix} 3 & 4 \\ 9 & 0 \end{pmatrix} + 6 \det \begin{pmatrix} 2 & 4 \\ 8 & 0 \end{pmatrix} - 7 \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} \\ &= (-5)(-36) + (6)(-32) - (7)(-6) \\ &= 180 - 192 + 42 = 30. \end{aligned}$$

La expansión por la tercera columna es

$$\begin{aligned} &4 \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 7 \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + 0 \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \\ &= (4)(-3) - (7)(-6) \\ &= -12 + 42 = 30. \end{aligned}$$

Las Ecs. 8.1 y 8.1F son útiles para calcular los determinantes, pero algunas de sus consecuencias teóricas son aún más importantes.

Dada la matriz A, consideremos la expresión

$$\sum_{i=1}^{n} (-1)^{i+j} x_i \det A_{ij},$$

donde $x_1, \ldots, x_n$ es un conjunto de n números cualesquiera. Usando 8.1, vemos que esto es un cierto determinante, de hecho, es la expansión por la columna j de la matriz obtenida a partir de A remplazando la columna j por $(x_1, \ldots, x_n)$. Ahora consideremos lo que sucede si tomamos $x_1, \ldots, x_n$ igual a los elementos $a_{1k}, a_{2k}, \ldots, a_{nk}$ de la columna k de A. Si $k = j$ tenemos, por supuesto, la expansión de det A por la columna j. Si $k \neq j$, tenemos el determinante de una matriz con dos columnas iguales (la j y la k) y, por el Teorema 7.8, el resultado es 0. Hemos demostrado

8.2 Teorema

Si A es una matriz n-por-n, entonces

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ik} \det A_{ij} = \begin{cases} \det A & \text{si } k = j \\ 0 & \text{si } k \neq j \end{cases}$$

Un argumento totalmente análogo, usando 8.1F en lugar de 8.1, de la forma «fila»:

8.2F Si A es una matriz n-por-n, entonces

$$\sum_{j=1}^{n}(-1)^{i+j}a_{kj}\det A_{ij} = \begin{cases}\det A & \text{si } k = i\\ 0 & \text{si } k \neq i\end{cases}$$

El número $(-1)^{i+j}\det A_{ij}$ se llama el **cofactor** ij de la matriz A. Lo abreviaremos por $\tilde{a}_{ij}$ y denotaremos por $\tilde{A}$ la matriz con elementos $\tilde{a}_{ij}$. El Teorema 8.2 puede ser formulado como una afirmación acerca del producto A^tA. El elemento j de este producto es el producto de la fila j de $\tilde{A}^t$ (es decir, la columna j de $\tilde{A}$) y la columna j de A. Es decir, es la suma $\sum_{i=j}^{n}\tilde{a}_{ij}a_{ik}$. Por el Teorema 8.2, esto es $\det A$ si $j = k$ y 0 en caso contrario. Por tanto, $\tilde{A}^tA$ es igual a $(\det A)I$, un múltiplo numérico de la matriz identidad. Un cálculo análogo usando 8.2 demuestra que $A\tilde{A}^t$ es también igual a $(\det A)I$. Si $\det A \neq 0$, podemos dividir $\tilde{A}^t$ por $\det A$ y obtener una matriz B tal que $AB = BA = I$. Por tanto, hemos demostrado

8.3 Teorema

Si $\det A \neq 0$, entonces A es invertible y $A^{-1} = (\det A)^{-1}\tilde{A}^t$, donde $\tilde{A}$ es la matriz de cofactores de A.

Una consecuencia importante del Teorema 8.3 es que si solo se sabe que una matriz A tiene un «inverso por la derecha» ($BA = I$) o un «inverso por la izquierda» ($AB = I$), entonces A es invertible y $A^{-1} = B$. Esto es válido puesto que aplicando la regla producto a $AB = I$ o $BA = I$, se tiene que

$$(\det A)(\det B) = \det I = 1.$$

Pero entonces, $\det A \neq 0$, de modo que A es invertible. Es claro que $A^{-1} = B$ (¿por qué?).

Ejemplo 2. Calcularemos el inverso de la matriz

$$A = \begin{pmatrix}2 & 3 & 4\\ 5 & 6 & 7\\ 8 & 9 & 0\end{pmatrix}$$

usada en el Ejemplo 1. Denotemos $\det A_{ij}$ por b_{ij}, entonces se calcula fácilmente que la matriz B es

$$\begin{pmatrix}-63 & -56 & -3\\ -36 & -32 & -6\\ -3 & -6 & -3\end{pmatrix}.$$

Para obtener la matriz de cofactores, introducimos los factores $(-1)^{i+j}$, lo que tiene el efecto de cambiar los signos de uno de cada dos elementos y obtenemos

$$\begin{pmatrix}-63 & 56 & -3\\ 36 & -32 & 6\\ -3 & 6 & -3\end{pmatrix}.$$

Finalmente, transponiendo y dividiendo por $\det A$ que, por el Ejemplo 1, se sabe es 30, tenemos que

$$A^{-1} = \tfrac{1}{30}\begin{pmatrix}-63 & 36 & -3\\ 56 & -32 & 6\\ -3 & 6 & -3\end{pmatrix},$$

como puede ser comprobado calculando AA^{-1} y $A^{-1}A$.

Un sistema de n ecuaciones lineales en n incógnitas

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1\\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2\\ &\vdots\\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n\end{aligned} \tag{1}$$

puede ser escrito en forma matricial como $A\mathbf{x} = \mathbf{b}$, donde A es la matriz n-por-n de coeficientes, cuyos elementos son a_{ij} y $\mathbf{x}$, $\mathbf{b}$ son vectores en $\mathcal{R}^n$. Efectuando la multiplicación matricial de la ecuación

$$\begin{pmatrix}a_{11} & \ldots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \ldots & a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} = \begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$$

muestra inmediatamente que esto es equivalente al sistema (1).

Si A es invertible, entonces $A\mathbf{x} = \mathbf{b}$ implica $A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$ o sea $\mathbf{x} = A^{-1}\mathbf{b}$. Por otra parte, $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b}$. En otras palabras, las ecuaciones tienen una solución única, $A^{-1}\mathbf{b}$. El elemento j del vector columna $A^{-1}\mathbf{b}$ es el producto matricial de la fila de A^{-1} y el vector $\mathbf{b}$. Si $\det A \neq 0$, podemos expresar los elementos de A^{-1} como cofactores de A y obtenemos

$$(\det A)^{-1}\sum_{i=1}^{n}(-1)^{i+j}(\det A_{ij})b_i$$

para este producto. Usando 8.1, vemos que esta expresión es simplemente $(\det A)^{-1}\det B^{(j)}$, donde $B^{(j)}$ se obtiene remplazando la columna j por A y por $\mathbf{b}$. Hemos demostrado:

8.4 Regla de Cramer

Si el determinante de la matriz de coeficiente de un sistema de n ecuaciones lineales con n incógnitas $x_1, \ldots, x_n$ es distinto de cero, entonces existe una solución única dada por

$$x_j = \frac{\det B^{(j)}}{\det A},$$

donde A es la matriz de coeficiente y $B^{(j)}$ se obtiene remplazando la columna j de A por la columna de números formada por los miembros de la derecha de las ecuaciones.

Ejemplo 3. Resolvamos el sistema

$$\begin{aligned} x_1 - 2x_2 + 4x_3 &= 1 \\ -x_1 + x_2 - x_3 &= 2 \\ 2x_1 + 3x_2 - x_3 &= 3. \end{aligned}$$

Tenemos

$$A = \begin{pmatrix} 1 & -2 & 4 \\ -1 & 1 & -1 \\ 2 & 3 & -1 \end{pmatrix}, \quad B^{(1)} = \begin{pmatrix} 1 & -2 & 4 \\ 2 & 1 & -1 \\ 3 & 3 & -1 \end{pmatrix},$$

$$B^{(2)} = \begin{pmatrix} 1 & 1 & 4 \\ -1 & 2 & -1 \\ 2 & 3 & -1 \end{pmatrix}, \quad B^{(3)} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & 2 \\ 2 & 3 & 3 \end{pmatrix}.$$

Expandiendo los determinantes por sus primeras filas, tenemos

$$\det A = (1)(2) - (-2)(3) + (4)(-5) = 2 + 6 - 20 = -12$$

$$\det B^{(1)} = (1)(2) - (-2)(1) + (4)(3) = 2 + 2 + 12 = 16$$

$$\det B^{(2)} = (1)(1) - (1)(3) + (4)(-7) = 1 - 3 - 28 = -30$$

$$\det B^{(3)} = (1)(-3) - (-2)(-7) + (1)(-5) = -3 - 14 - 5 = -22.$$

Entonces $x_1 = -\frac{16}{12} = -\frac{4}{3}$, $x_2 = \frac{30}{12} = \frac{5}{2}$, $x_3 = \frac{22}{12} = \frac{11}{6}$.

En esta sección no hemos hecho ninguna afirmación acerca de lo que sucede si el determinante de la matriz de coeficiente de un sistema de ecuaciones es cero. Una consecuencia fácil de la regla producto es que una matriz con determinante cero no puede ser invertible (véase el Ejercicio 1 de la Sec. 7), además se demostrará en el Cap. 2, donde discutiremos en detalle la solución de sistemas de ecuaciones lineales, que en este caso el sistema no tiene solución o tiene infinitas soluciones. A pesar de que la regla de Cramer y la fórmula para A^{-1}

en términos de los cofactores son de gran importancia como resultados teóricos, y son bastante útiles para resolver sistemas de dos y tres ecuaciones lineales, ellos son menos eficientes para sistemas más grandes que los métodos de la Sec. 1 del Cap. 2.

Como 8.1 con $j = 1$ es exactamente como 7.1, excepto que se refiere a la primera columna en lugar de la primera fila de una matriz, es claro que, si transponemos una matriz (lo cual simplemente intercambia los papeles de filas y columnas) no se deberá alterar el valor del determinante. La afirmación formal y su demostración son dadas a continuación.

8.5 Teorema

Para toda matriz cuadrada A, $\det A^t = \det A$.

Demostración. Sea $B = A^t$. Por la definición de transpuesta, $b_{ij} = a_{ji}$ y es fácil ver que $B_{ij} = A^t_{ji}$. Por tanto, por definición

$$\begin{aligned} \det B &= b_{11} \det B_{11} - b_{12} \det B_{12} + \ldots + (-1)^{n+1} b_{1n} \det B_{1n} \\ &= a_{11} \det A^t_{11} - a_{21} \det A^t_{21} + \ldots + (-1)^{n+1} a_{n1} \det A^t_{n1}. \end{aligned}$$

Las A_{ij} son matrices $(n-1)$-por-$(n-1)$ y, por hipótesis de inducción, podemos remplazar $\det A^t_{i1}$ por $\det A_{i1}$ en la fórmula anterior. Usando 8.1 (con $j = 1$), vemos que el resultado es $\det A$ y hemos demostrado que $\det A^t = \det B = \det A$. Ahora podemos tomar cualquier teorema acerca de determinantes que esté relacionado con las columnas de una matriz y obtener de inmediato el teorema correspondiente para las filas de una matriz, simplemente aplicando el teorema recién demostrado a las transpuestas. Generalmente, no nos daremos el trabajo de hacer explícitos estos teoremas. Nos referiremos a la versión «fila» de un resultado numerado con el mismo número con una F después. (Ya hemos designado por 8.1F y 8.2F de acuerdo con esta convención.) Por ejemplo, el Teorema 7.2F es la afirmación: Si B se obtiene de A multiplicando una fila por un número r, entonces $\det B = r \det A$.

El Teorema 8.5 implica que, a partir de cualquier teorema acerca de determinantes que tenga que ver con las columnas de matrices, se puede deducir el teorema correspondiente para las filas aplicando el teorema dado a las transpuestas de las matrices. En particular, 8.1F es una consecuencia de 8.1 y 8.5. Varios de los teoremas de la Sec. 7 tienen versiones «fila», por ejemplo, de 7.4 se deduce 7.2F. Si B se obtiene de A por multiplicación de una fila por un número r, entonces $\det B = r \det A$. No nos molestaremos en escribir las versiones fila de las otras afirmaciones, pero podremos referirnos a ellas mediante la afirmación original seguida de una F. En la sección anterior se demostró el resultado importante que el determinante es una función lineal de cada una de sus columnas. Ahora tenemos que el determinante también es función lineal de cada una de sus filas.

El teorema siguiente (y el correspondiente para filas) nos da una manera eficiente para calcular el determinante de una matriz grande.

8.6 Teorema

Si C y A son matrices n-por-n y C se obtiene de A sumando un múltiplo numérico de una columna a otra, entonces $\det C = \det A$.

Demostración. Supongamos que C es igual a A, excepto que $\mathbf{c}_j = \mathbf{a}_j + r\mathbf{a}_i$. Sea la matriz obtenida de A remplazando $\mathbf{a}_j$ por $r\mathbf{a}_i$. Por el Teorema 7.4, $\det C = \det A + \det B$. Por el Teorema 7.2, $\det B$ es r veces el determinante de una matriz con dos columnas iguales. Por tanto, $\det B = 0$ y $\det C = \det A$.

Ejemplo 4. (a) Sean

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & 5 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 3 & 0 \\ 2 & -4 & 5 \\ 3 & 5 & 4 \end{pmatrix}.$$

La tercera columna de C es igual a la tercera columna de A más 2 veces la primera columna. Del Ejemplo 5, Sec. 7, $\det A = -20$, de donde $\det C = -20$. Esto se puede comprobar calculando

$$\begin{aligned} \det C &= (1)(-16 - 25) - (3)(8 - 15) + (0)\big(10 - (-12)\big) \\ &= -41 + 21 + 0 = -20. \end{aligned}$$

(b) Sea

$$A = \begin{pmatrix} 2 & 4 & -1 & 0 \\ 3 & 0 & 2 & 3 \\ -1 & 2 & 3 & 1 \\ 0 & 1 & -2 & -1 \end{pmatrix}$$

Sumando 2 veces la tercera columna a la primera y 4 veces la tercera columna a la segunda, obtenemos

$$B = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 7 & 8 & 2 & 3 \\ 5 & 14 & 3 & 1 \\ -4 & -7 & -2 & -1 \end{pmatrix},$$

por el Teorema 8.6, $\det A = \det B$. La expresión de B tiene un solo término distinto de cero y obtenemos

$$\det B = (-1) \det \begin{pmatrix} 7 & 8 & 3 \\ 5 & 14 & 1 \\ -4 & -7 & -1 \end{pmatrix}$$

$$= -\det \begin{pmatrix} 7 & 1 & 3 \\ 5 & 9 & 1 \\ -4 & -3 & -1 \end{pmatrix}$$ [restar la columna 1 a la columna 2]

$$= -\det \begin{pmatrix} 0 & 1 & 0 \\ -58 & 9 & -26 \\ 17 & -3 & 8 \end{pmatrix}$$ [restar 7 veces la columna 2 a la columna 1 y 8 veces la columna 2 a la columna 3].

Entonces, $\det B = -(-1)\big((-58)(8) - (17)(-26)\big) = -22$.

Este último ejemplo ilustra una manera de calcular determinantes que es, generalmente, mucho más eficiente que el método de las Ecs. 8.1 a 8.1F, en el caso de matrices más grandes que 3-por-3.

El lema siguiente es un paso preliminar para la demostración de 8.1. Recordamos que $\mathbf{e}_i$ es el i-ésimo vector de la base natural de $\mathcal{R}^n$ y sus coordenadas son 0, excepto la i-ésima, que es 1.

8.7 Lema

Si la primera columna de la matriz A es $\mathbf{e}_i$, entonces

$$\det A = (-1)^{i+1} \det A_{i1}.$$

Demostración. Por la Definición 7.1,

$$\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det A_{1j}.$$

Si la primera columna de A es $\mathbf{e}_1$, entonces $a_{11} = 1$, mientras que, para $j > 1$, la primera columna de A_{1j}, de modo que $\det A_{1j} = 0$ por el Teorema 7.3. Por tanto, la expresión para $\det A$ se reduce a un solo término, $\det A_{11}$ y el Lema 8.7 queda demostrado en el caso $i = 1$. Para $i > 1$, debemos usar como hipótesis de inducción que el Lema 8.7 es válido para matrices $(n-1)$-por-$(n-1)$. En este caso, $a_{11} = 0$ y

$$\det A = \sum_{j=2}^{n} (-1)^{j+1} a_{1j} \det A_{1j}.$$

Cada menor A_{1j} tiene $\mathbf{e}_{i-1}$ como primera columna. (Si quitamos la primera coordenada al vector $\mathbf{e}_i$ de $\mathcal{R}^n$, obtenemos el vector $\mathbf{e}_{i-1}$ de $\mathcal{R}^{n-1}$.) Por la hipótesis de inducción,

$$\det A_{1j} = (-1)^i \det A_{1j,i1},$$

donde hemos denotado por $A_{1j,i1}$ a la matriz obtenida de A quitando la primera fila, la columna j, la fila i y la primera columna. Sea $B = A_{i1}$. Entonces (puesto que la primera columna de B se obtiene de la segunda columna de A, etc.),

$A_{1j,i1} = B_{1,j-1}$ y $a_{1j} = b_{1,j-1}$. Combinando las ecuaciones obtenidas, tenemos

$$\det A = \sum_{j=2}^{n} (-1)^{j+1} b_{1,j-1} (-1)^{i} \det B_{1,j-1}.$$

De la Ec. 7.1 se tiene que el miembro de la derecha de esta ecuación es $(-1)^{i+1} \det B$, que es igual a $(-1)^{i+1} \det A_{i1}$, como se quería demostrar.

Demostración de 8.1. Primero demostraremos 8.1 para el caso especial en que $j = 1$. La primera columna de A puede ser escrita en la forma $\mathbf{a}_1 = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \ldots + a_{1n}\mathbf{e}_n$, de modo que, por el Teorema 8.5 y el Lema 8.7,

$$\begin{aligned} \det A &= \det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \\ &= a_{11} \det(\mathbf{e}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + a_{12} \det(\mathbf{e}_2, \mathbf{a}_2, \ldots, \mathbf{a}_n) \\ &\quad + a_{1n} \det(\mathbf{e}_n, \mathbf{a}_2, \ldots, \mathbf{a}_n) \\ &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}. \end{aligned}$$

Para demostrar el caso general, consideramos la matriz B obtenida a partir de A moviendo la columna j a la primera posición. Este movimiento involucra una serie de $j-1$ intercambios de columnas adyacentes, de modo que, por el Teorema 6.9, $\det A = (-1)^{j-1} \det B$. Para todo i, tenemos que $b_{i1} = a_{ij}$ y $B_{i1} = A_{ij}$. Por tanto,

$$\begin{aligned} \det A &= (-1)^{j-1} \det B \\ &= (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i+1} b_{i1} \det B_{i1} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}, \end{aligned}$$

como se quería demostrar.

EJERCICIOS

1. Usando los casos adecuados de 8.1 u 8.1F, exprese cada una de las expresiones siguientes como función lineal de las x.

(a) $\begin{pmatrix} 0 & 3 & 6 \\ -2 & -1 & 6 \\ x_1 & x_2 & x_3 \end{pmatrix}$.

(b) $\begin{pmatrix} -2 & 3 & x_1 & 1 \\ 0 & -1 & x_2 & 0 \\ -1 & 2 & x_3 & 1 \\ 1 & 0 & x_4 & 0 \end{pmatrix}$.

2. Demuestre que el producto $P_1 P_2 \ldots P_k$ de matrices cuadradas es invertible si, y solo si, cada una de ellas lo es. Por tanto demuestre que si A y B son matrices cuadradas y $AB = I$, entonces A y B son invertibles y son inversas una de la otra.

3. Encuentre las inversas de

(a) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 0 & -1 & 3 \\ 2 & 5 & -4 \\ -3 & 7 & 1 \end{pmatrix}$,

(c) $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 \\ -1 & 0 & 0 & 2 \\ -3 & 0 & -2 & 0 \end{pmatrix}$.

[*Resp.* (b) $\frac{1}{77}\begin{pmatrix} 33 & 22 & -11 \\ 10 & 9 & 6 \\ 29 & 3 & 2 \end{pmatrix}$.]

4. Resuelva los sistemas

(a) $7x + 6y = 5$
$6x + 5y = -3$.

(b) $2x + y = 0$
$3y + z = 1$
$4z + x = 2$.

[*Resp.* $x = -\frac{2}{25}, y = \frac{4}{25}, z = \frac{13}{25}$.]

(c) $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= -1 \\ x_1 - x_2 + 2x_4 &= 0 \\ 3x_2 - x_3 &= 3 \\ x_2 - x_4 &= 0. \end{aligned}$

5. Use el método del Ejemplo 4(b) para calcular

(a) $\det \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 \\ 1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \end{pmatrix}$.

[*Resp.* (a) 32.]

(b) $\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix}$.

6. (a) Calcule

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 5 & 6 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

(b) Una matriz A como la de la parte (a) de la cual todos los elementos debajo de la diagonal principal son cero se llama **triangular.** Demuestre que si A es una matriz triangular entonces det A es igual al producto de los elementos diagonales.

7. (a) Una matriz de la forma

$$A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & 0 \\ c_1 & a_2 & b_2 & 0 & 0 \\ 0 & c_2 & a_3 & b_3 & 0 \\ 0 & 0 & c_3 & a_4 & b_4 \\ 0 & 0 & 0 & c_4 & a_5 \end{pmatrix}$$

que es cero excepto por los elementos en o adyacentes a la diagonal se llama matriz **tridiagonal.** Sea d_k el determinante del menor k-por-k formado por las primeras k columnas y filas de A, de modo que, por ejemplo, $d_1 = a_1$ y $d_2 = a_1a_2 = b_1c_1$. Demuestre que para $k \geq 3$, $d_k = a_kd_{k-1} - b_{k-1}c_{k-1}d_{k-2}$.

(b) Considere matrices tridiagonales tales que todos los elementos en la diagonal tienen el valor 2 y que los adyacentes a la diagonal tienen el valor 1. Sea d_n el determinante de una matriz n-por-n de este tipo. Encuentre una fórmula para d_n. [*Sugerencia*: Empiece considerando lo que sucede para $n = 2, 3, 4$.]

2

Algebra lineal

1. Ecuaciones lineales, matrices inversas

Un **sistema de ecuaciones lineales** es un conjunto finito de ecuaciones

1.1
$$\begin{array}{c} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{array}$$

donde las a y las b están dadas y las x deben ser determinadas. El sistema completo puede ser escrito en forma matricial como $A\mathbf{x} = \mathbf{b}$, donde A es la matriz m-por-n de coeficientes con elementos a_{ij}, $\mathbf{b}$ es un vector columna en $\mathcal{R}^m$ y $\mathbf{x}$ es un vector columna en $\mathcal{R}^n$. Efectuando la multiplicación matricial en la ecuación

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

se ve de inmediato que esto es equivalente al sistema 1.1.

A menudo sucede en las aplicaciones que hay el mismo número de ecuaciones y de incógnitas, de modo que $m = n$. Entonces, si el determinante de la matriz de coeficientes no es cero, la solución puede obtenerse usando la regla de Cramer de la Sec. 8, del Cap. 1. En esta sección usamos métodos que no dependen de los determinantes y no suponemos que $m = n$, cuando n es mayor que 3, estos métodos son más eficientes que la regla de Cramer.

Cualquier vector $\mathbf{c}$ de $\mathbb{C}^n$ tal que $A\mathbf{c} = \mathbf{b}$ es *una* solución del sistema. Como demostraremos, algunos sistemas no tienen solución, otros tienen exactamente una y otros infinitas soluciones.

Decimos que dos sistemas son **equivalentes** si ellos tienen el mismo conjunto de soluciones. Nuestro procedimiento consistirá en tomar un sistema dado y alterarlo en una sucesión de pasos para obtener un sistema equivalente para el cual las soluciones son obvias. Ilustraremos este proceso con un ejemplo antes de dar la descripción general.

Ejemplo 1.

$$\begin{aligned} 3x + 12y + 9z &= 3 \\ 2x + 5y + 4z &= 4 \\ -x + 3y + 2z &= -5. \end{aligned}$$

Multiplicamos la primera ecuación por $\frac{1}{3}$, de modo que el coeficiente de x se hace igual a 1 y se tiene

$$\begin{aligned} x + 4y + 3z &= 1 \\ 2x + 5y + 4z &= 4 \\ -x + 3y + 2z &= -5. \end{aligned}$$

Sumamos (-2) veces la primera ecuación a la segunda y remplazamos la segunda ecuación por el resultado. Esto hace que el coeficiente de x en la segunda ecuación sea 0. Se tiene entonces

$$\begin{aligned} x + 4y + 3z &= 1 \\ -3y - 2z &= 2 \\ -x + 3y + 2z &= -5. \end{aligned}$$

Sumamos (-2) veces la primera ecuación a la segunda y remplazamos la segunda ecuación por el resultado. Entonces el coeficiente de x en la segunda ecuación es 0 y tenemos:

$$\begin{aligned} x + 4y + 3z &= 1 \\ -3y - 2z &= 2 \\ 7y + 5z &= -4. \end{aligned}$$

Multiplicamos la segunda ecuación por $-\frac{1}{3}$ y obtenemos

$$\begin{aligned} x + 4y + 3z &= 1 \\ y + \tfrac{2}{3}z &= -\tfrac{2}{3} \\ 7y + 5z &= -4. \end{aligned}$$

Sumamos (-4) veces la segunda ecuación a la primera y (-7) veces la segunda a la tercera, obteniendo así

$$\begin{aligned} x + \tfrac{1}{3}z &= \tfrac{11}{3} \\ y + \tfrac{2}{3}z &= -\tfrac{2}{3} \\ \tfrac{1}{3}z &= \tfrac{2}{3}. \end{aligned}$$

Multiplicamos la tercera ecuación por 3 y tenemos

$$\begin{aligned} x + \tfrac{1}{3}z &= \tfrac{11}{3} \\ y + \tfrac{2}{3}z &= -\tfrac{2}{3} \\ z &= 2. \end{aligned}$$

Sumamos $(-\frac{1}{3})$ veces la tercera ecuación a la primera y $(-\frac{2}{3})$ veces la tercera a la ecuación, obteniendo

$$\begin{aligned} x &= 3 \\ y &= -2 \\ z &= 2. \end{aligned}$$

Es evidente que este sistema tiene una sola solución: el vector columna

$$\begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}.$$

Es fácil verificar, por sustitución en la ecuación original, que hemos encontrado una solución. Por supuesto, esta verificación no elimina la posibilidad que las ecuaciones originales puedan tener otras soluciones. De hecho, el sistema final es equivalente al sistema original y tiene el mismo conjunto de soluciones. Lo mismo es verdad para cualquier par de sistemas tales que uno se puede obtener del otro por operaciones del tipo ilustrado en este ejemplo. Antes de demostrar esto, debemos expresar exactamente cuáles operaciones son permisibles e investigar sus propiedades.

Las operaciones usadas fueron «multiplicar una ecuación por un número» y «sumar a una ecuación un múltiplo de otra». Preferimos dar las definiciones formales en términos de matrices y, por tanto, considerar la ecuación matricial $A\mathbf{x} = \mathbf{b}$.

Definimos tres tipos de **operaciones elementales** que pueden ser aplicadas a una matriz M cualquiera:

Una **multiplicación elemental** remplaza una fila de M por un múltiplo numérico de la misma fila, el multiplicador debe ser no cero.

Una **modificación elemental** remplaza una fila de M por la suma de esa fila y un múltiplo numérico de otra fila.

Una **transposición elemental** intercambia dos filas de M. En el Ejemplo 1, no usamos transposiciones, pero ellas son útiles en algunas ecuaciones.

Es importante entender que si se aplica una operación elemental a ambos miembros de una ecuación $A\mathbf{x} = \mathbf{b}$, el resultado es una nueva ecuación matricial que es satisfecha por todo vector $\mathbf{x}$ que satisface la ecuación original. Igualmente importante es el hecho de que toda operación elemental tiene asociado una **operación elemental inversa** mediante la cual se puede anular el cambio efectuado por la operación original. Por ejemplo, la multiplicación de una fila por un número $r \neq 0$ se «anula» multiplicando la misma fila por $1/r$. Análogamente, una modificación elemental se anula sustrayendo el mismo múltiplo numérico de la misma fila en lugar de sumarlo. Por último, una transposición se anula intercambiando nuevamente las filas.

El sistema de ecuaciones del Ejemplo 1 tiene la siguiente forma como ecuación matricial

$$\begin{pmatrix} 3 & 12 & 9 \\ 2 & 5 & 4 \\ -1 & 3 & 2 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}.$$

Si aplicamos la operación elemental de multiplicar la primera fila por $\frac{1}{3}$ a las matrices de ambos términos, obtenemos

$$\begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 4 \\ -1 & 3 & 2 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix},$$

que es la forma matricial del sistema de ecuaciones en el segundo paso del Ejemplo 1. De la misma manera, cada operación hecha al sistema de ecuaciones es equivalente a efectuar una operación elemental a las matrices en ambos lados de la ecuación matricial equivalente. En la última etapa, la ecuación matricial tiene la forma

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}$$

la cual, puesto que la matriz de la izquierda es la matriz identidad, significa simplemente que $\mathbf{x}$ es igual al vector de la derecha.

El teorema que justifica nuestro método para resolver ecuaciones lineales es el siguiente.

1.2 Teorema

Si el sistema $A_1\mathbf{x} = \mathbf{b}_1$ se transforma en el sistema $A_2\mathbf{x} = \mathbf{b}_2$ mediante la aplicación de una sucesión de operaciones a A_1 para obtener A_2 y la *misma* sucesión de operaciones elementales a $\mathbf{b}_1$ para obtener $\mathbf{b}_2$, entonces los dos sistemas son equivalentes en el sentido de que tienen las mismas soluciones.

Demostración. Si $\mathbf{x}$ es una solución de $A_1\mathbf{x} = \mathbf{b}_1$, entonces, aplicando una operación elemental a ambos miembros, se obtiene una nueva ecuación matricial que también es satisfecha por $\mathbf{x}$. Por tanto, $\mathbf{x}$ es solución de $A_2\mathbf{x} = \mathbf{b}_2$, ya que esta ecuación se obtiene mediante una sucesión finita de operaciones de este tipo. Conversamente, si hemos transformado $A_1\mathbf{x} = \mathbf{b}_1$ en $A_2\mathbf{x}_2 = \mathbf{b}_2$ mediante una sucesión de operaciones elementales podemos encontrar una sucesión de operaciones elementales inversas que transforma la segunda ecuación en la primera. El mismo argumento usado en la primera parte de la demostración muestra que toda solución de $A_2\mathbf{x} = \mathbf{b}_2$ es solución de $A_1\mathbf{x} = \mathbf{b}_1$.

Ejemplo 2. Daremos un ejemplo de un sistema de ecuaciones con infinitas soluciones. Consideremos la ecuación matricial

$$\begin{pmatrix} 1 & -2 & -3 \\ \frac{1}{2} & -2 & -\frac{3}{2} \\ -3 & 5 & 4 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 2 \\ 7 \\ 0 \end{pmatrix}.$$

Sumamos $(-\frac{1}{2})$ veces la primera fila a la segunda y después 3 veces la primera fila a la tercera fila para obtener ceros en el segundo y tercer elemento de la primera columna, es decir, obtenemos

$$\begin{pmatrix} 1 & -2 & -3 \\ 0 & -1 & -5 \\ 0 & -1 & -5 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 2 \\ 6 \\ 6 \end{pmatrix}.$$

Multiplicamos la segunda fila por (-1) para obtener

$$\begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 5 \\ 0 & -1 & -5 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 2 \\ -6 \\ 6 \end{pmatrix}.$$

Sumamos 2 veces la segunda fila a la primera y después sumamos la segunda fila a la tercera para obtener

$$\begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{x} = \begin{pmatrix} -10 \\ -6 \\ 0 \end{pmatrix}.$$

En la etapa correspondiente del Ejemplo 1, efectuamos una multiplicación elemental para hacer igual a 1 el elemento de la tercera fila. Como tenemos una fila de ceros, no podemos hacer lo mismo. Escribamos $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, y tra-

duzcamos la ecuación matricial a un sistema de ecuaciones lineales. El resultado es

$$\begin{aligned} x \qquad + 7z &= -10 \\ y + 5z &= -6 \\ 0x + 0y + 0z &= \quad 0. \end{aligned}$$

La tercera ecuación se satisface para todo valor de x, y, z. Las primeras dos ecuaciones pueden ser reescritas en la forma $x = -10 - 7z$ y $y = -6 - 5z$. Por tanto, para cualquier valor de z,

$$\begin{pmatrix} -7z - 10 \\ -5z - 6 \\ z \end{pmatrix} = z\begin{pmatrix} -7 \\ -5 \\ 1 \end{pmatrix} + \begin{pmatrix} -10 \\ -6 \\ 0 \end{pmatrix}$$

es una solución y toda solución tiene esta forma para algún valor de z. De esta manera, hemos descrito el conjunto de soluciones de

$$\begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{x} = \begin{pmatrix} -10 \\ -6 \\ 0 \end{pmatrix},$$

y, por el Teorema 1.2, sabemos que éste es el conjunto de soluciones de la ecuación matricial original.

Ejemplo 3. Consideremos la ecuación matricial

$$\begin{pmatrix} 1 & -2 & -3 \\ \frac{1}{2} & -2 & -\frac{3}{2} \\ -3 & 5 & 4 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 2 \\ 7 \\ 2 \end{pmatrix}.$$

La matriz de la izquierda es la misma que la del Ejemplo 2. Si efectuamos la misma sucesión de operaciones elementales obtenemos

$$\begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{x} = \begin{pmatrix} -10 \\ -6 \\ 2 \end{pmatrix}.$$

Para todo $\mathbf{x}$, la tercera fila del producto $\begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}$ es cero, ya que la tercera fila del factor de la izquierda es cero. Por tanto, ningún valor de $\mathbf{x}$ puede dar un vector columna con 2 en la tercera fila, y concluimos que la ecuación *no* tiene solución.

(Si hacemos $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ y escribimos la ecuación matricial como un sistema de ecuaciones, tenemos

$$\begin{aligned} x \qquad + 7z &= -10 \\ y + 5z &= \quad -6 \\ 0x + 0y + 0z &= \quad 2. \end{aligned}$$

Evidentemente, la última ecuación no puede ser satisfecha para ningún valor de x, y, z.)

En estos ejemplos hemos usado operaciones elementales para transformar el sistema original de ecuaciones en uno equivalente cuyas soluciones son muy fáciles de encontrar. La propiedad del conjunto final de ecuaciones que hace obvio encontrar las soluciones es que cada ecuación involucra una variable que no aparece en las otras ecuaciones. Así, en el Ejemplo 2, encontramos $x = -10 - 7z$, $y = -6 - 5z$ y, como la primera ecuación contiene a z y no a y, y la segunda contiene a y y no a x, podemos encontrar todos los valores de (x, y, z) que satisfacen ambas ecuaciones considerando cada ecuación por separado. Las definiciones que siguen expresan esta «no intervención» en términos de la matriz de coeficientes. En la práctica, no hay dificultad (más allá del trabajo de hacer las operaciones aritméticas) en reducir cualquier sistema de ecuaciones a un sistema que tiene esta propiedad de «no intervención». Para un sistema reducido de este tipo, es obvio que no existe una solución, como en el Ejemplo 3, o bien la solución o el conjunto de soluciones puede ser descrito explícitamente como en los Ejemplos 1 y 2. En la Sec. 3 daremos la demostración de que siempre se puede usar este método.

Supongamos que tenemos una matriz cuadrada A y podemos convertirla en la matriz identidad mediante una sucesión de operaciones elementales, tal como lo hiciéramos en el Ejemplo 1. Entonces el Teorema 1.2 afirma que si le aplicamos la misma sucesión de operaciones al vector $\mathbf{b}$ y obtenemos $\mathbf{c}$ como resultado, entonces $A\mathbf{x} = \mathbf{b}$ si, y solo si, $I\mathbf{x} = \mathbf{c}$. En otras palabras, $\mathbf{c}$ es la solución (única) de $A\mathbf{x} = \mathbf{b}$. Aplicando las mismas operaciones a la matriz identidad se obtiene una matriz C cuya columna j, $\mathbf{c}_j$, es la solución de $A\mathbf{c}_j = \mathbf{e}_j$, donde $\mathbf{e}_j$ es la columna j de I. Por tanto, $AC = I$. Esto sugiere que $C = A^{-1}$ y que hemos encontrado un método para calcular matrices inversas. Esto es efectivamente así; formalmente tenemos

1.3 Teorema

Si una matriz n-por-A puede ser transformada en la matriz identidad n-por-n, I, mediante una sucesión de operaciones elementales entonces A es invertible y A^{-1} es igual al resultado de la aplicación a I de la misma sucesión de operaciones.

Demostración. La discusión anterior muestra que, si C es el resultado de la aplicación a I de las mismas operaciones elementales que transforman A en I, entonces $AC = I$. Para demostrar que A es invertible y que $C = A^{-1}$, debemos demostrar que $CA = I$. El argumento está basado en el hecho de que las operaciones elementales son reversibles. Ahora, como C fue construida aplicando operaciones elementales a I, se tiene que C puede ser transformada en I mediante operaciones elementales. Si aplicamos a C el mismo argumento que aplicáramos a A, se deduce que existe una matriz D tal que $CD = I$. Entonces

$$A = AI = A(CD) = (AC)D = ID = D;$$

de modo que $D = A$. Por tanto, $CA = I$, como se quería demostrar.

Ejemplo 4. Encontraremos la inversa de

$$A = \begin{pmatrix} 2 & 4 & 8 \\ 1 & 0 & 0 \\ 1 & -3 & -7 \end{pmatrix}.$$

Empezamos con

$$\begin{pmatrix} 2 & 4 & 8 \\ 1 & 0 & 0 \\ 1 & -3 & -7 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Sumamos -2 veces la segunda fila a la primera y -1 veces la segunda fila a la primera y obtenemos

$$\begin{pmatrix} 0 & 4 & 8 \\ 1 & 0 & 0 \\ 0 & -3 & -7 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Multiplicamos la primera fila por $\frac{1}{4}$ y después sumamos 3 veces la primera fila a la tercera, obteniendo

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{4} & -\frac{1}{2} & 0 \\ 0 & 1 & 0 \\ \frac{3}{4} & -\frac{5}{2} & 1 \end{pmatrix}.$$

Multiplicamos la tercera fila por -1 y después sumamos -2 veces la tercera fila a la primera, obteniendo

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \frac{7}{4} & -\frac{11}{2} & 2 \\ 0 & 1 & 0 \\ -\frac{3}{4} & \frac{5}{2} & -1 \end{pmatrix}.$$

Transponiendo la primera y segunda filas, obtenemos

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ \frac{7}{4} & -\frac{11}{2} & 2 \\ -\frac{3}{4} & \frac{5}{2} & -1 \end{pmatrix}.$$

La matriz de la derecha es A^{-1}, como puede demostrarse directamente multiplicándola por A.

EJERCICIOS

1. Resuelva los siguientes sistemas de ecuaciones

(a) $$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -4 \\ 1 & 3 & 9 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}.$$

(b) $x_1 + x_2 = 1$
$x_2 + x_3 = 2$
$x_3 + x_4 = 3$
$x_4 + x_1 = 4.$

(c) $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 10 \\ 11 \\ 12 \end{pmatrix}.$$

2. Demuestre que el Teorema 1.2 implica el siguiente resultado más general. Si una operación elemental transforma la matriz A_1 en A_2 y la misma operación transforma B_1 en B_2 entonces el conjunto de las matrices X tales que $A_1X = B_1$ es igual al conjunto de las matrices X tales que $A_2X = B_2$.

3. Resuelva las ecuaciones matriciales

(a) $$\begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 & -3 & 4 \\ 1 & 2 & 0 \end{pmatrix}.$$

(b) $$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 0 \\ -4 & 1 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 & 1 \\ -5 & 4 \\ 0 & -3 \end{pmatrix}.$$

4. Para cada una de las matrices siguientes, encuentre A^{-1} si A es invertible y encuentre una solución distinta de cero de $A\mathbf{x} = 0$ si A no es invertible.

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 5 \\ -2 & 0 & 1 \end{pmatrix}$ [*Resp.* $\begin{pmatrix} 1 & 0 & 0 \\ -13 & 1 & -5 \\ 2 & 0 & 1 \end{pmatrix}$.]

(b) $\begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 0 & 3 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & -1 & 0 & 0 \\ 0 & 5 & -2 & 0 \\ 0 & 0 & 6 & -3 \\ 0 & 0 & 0 & 7 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 5 & 4 & 2 \\ -1 & -3 & 1 & 0 \\ 0 & 3 & 7 & 5 \end{pmatrix}$

2. Algunas aplicaciones

En esta sección veremos algunas aplicaciones de los vectores y las ecuaciones lineales. La selección de los ejemplos ha sido hecha de modo que se eviten las complicaciones técnicas de los campos de aplicación. Varios de los ejemplos que consideraremos están relacionados con la noción de **red** que definimos como una colección finita de puntos, o nodos, algunos de los cuales pueden estar unidos a otros por segmentos de rectas. Teóricamente, no es importante si visualizamos la red en el espacio de dos o de tres dimensiones, elegimos el que sea pictóricamente más conveniente. La Fig. 1 muestra algunas redes.

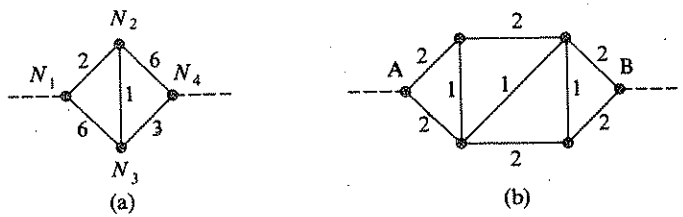


Figura 1

Ejemplo 1. Algunos circuitos eléctricos pueden ser considerados como redes en las cuales cada línea representa un alambre con una resistencia determinada. Cuando se conecta el circuito a una batería u otra fuente de poder, la corriente fluye en las conexiones y se establece un valor del potencial eléctrico en cada nodo. El flujo de corriente a través de muchos tipos de alambre es proporcional a la diferencia entre los potenciales de sus extremos. En una fórmula,

$$c_{ij} = \frac{(v_i - v_j)}{r_{ij}} \tag{1}$$

donde c_{ij} es la corriente que fluye del nodo i al nodo j, r_{ij} es la resistencia de la conexión entre los nodos i y j y v_i y v_j son los valores de los potenciales eléctricos en los nodos i y j. (Las unidades apropiadas de medida son los amperes para la corriente, los ohms para la resistencia y los voltios para el potencial.) Un valor negativo de la corriente de i a j se interpreta como una corriente de j a i.

La Fig. 1(a) muestra un circuito con cuatro nodos y cinco segmentos con la resistencia en cada uno de ellos indicado a su lado. Supongamos que se conecta una fuente de poder a los nodos 1 y 4 de modo que se mantengan los valores $v_1 = 12$ y $v_4 = 0$. Como el nodo 2 no tiene conexión externa, la corriente que fluye hacia él debe equilibrarse con la que sale, de modo que si se toman en cuenta los signos, la suma de las corrientes en el nodo 2 debe ser cero. Usando (1), obtenemos la ecuación

$$\tfrac{1}{2}(v_2 - v_1) + (v_2 - v_3) + \tfrac{1}{6}(v_2 - v_4) = 0.$$

Si la reescribimos en la forma

$$(\tfrac{1}{2} + 1 + \tfrac{1}{6})v_2 = \tfrac{1}{2}v_1 + v_3 + \tfrac{1}{6}v_4,$$

vemos que v_2 es el término medio de v_1, v_3 y v_4, con coeficientes que son los inversos de las resistencias en las líneas que unen al nodo 2 con los demás nodos. En cada nodo que no tenga una conexión externa se tendrá una ecuación similar. Así, en el nodo 3 tendremos

$$(\tfrac{1}{6} + 1 + \tfrac{1}{3})v_3 = \tfrac{1}{6}v_1 + v_2 + \tfrac{1}{3}v_4.$$

Si introducimos los valores $v_1 = 12$ y $v_4 = 0$ y reordenamos los términos, tenemos

$$\tfrac{5}{3}v_2 - v_3 = 6$$
$$-v_2 + \tfrac{3}{2}v_3 = 2.$$

Si resolvemos este sistema tenemos que $v_2 = \frac{22}{3} \approx 7{,}33$ y $v_3 = \frac{56}{9} \approx 6{,}22$. Una vez conocidos los potenciales, las otras cantidades se obtienen fácilmente. Por ejemplo, la corriente del nodo 1 al 2 es $(v_1 - v_2)/r_{12} = \frac{1}{2}(12 - 7{,}33) = 2{,}33$. Análogamente, la corriente del nodo 1 al 3 es $\frac{1}{16}(12 - 6{,}22) = 0{,}96$. La corriente total que fluye del nodo 1 al resto del circuito es $2{,}33 + 0{,}96 = 3{,}29$, que debe ser igual a la corriente que llega al nodo 1 de la fuente externa.

Ejemplo 2. Podemos considerar que vectores en $\mathcal{R}^2$ o $\mathcal{R}^3$ representan fuerzas que actúan en un punto que suponemos por conveniencia es el origen. La dirección de la flecha es la dirección en que actúa la fuerza y su largo es la magnitud de la fuerza. Nuestra suposición física fundamental es que si varias fuerzas actúan sobre un punto, entonces la **fuerza resultante** que actúa sobre el punto está representada por la suma de los vectores fuerza que actúan sobre el punto. En la Fig. 2, tenemos dos dibujos distintos. La flecha resultante $\mathbf{r}$ aparece solo en la Fig. 2(a). En términos numéricos, supongamos que los vectores fuerza de la Fig. 2(a) están en el plano $\mathcal{R}^2$ con el origen como punto de acción. Si tenemos

$$\mathbf{f}_1 = (-1, 3), \qquad \mathbf{f}_2 = (4, 3), \qquad \mathbf{f}_3 = (-2, -4), \tag{2}$$

entonces, por definición,

$$\mathbf{r} = \mathbf{f}_1 + \mathbf{f}_2 + \mathbf{f}_3 = (1, 2).$$

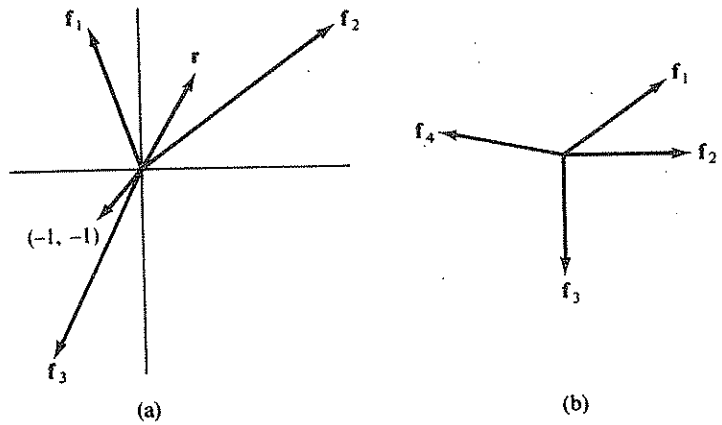


Figura 2

Supongamos ahora que conocemos las direcciones de los tres vectores fuerza y queremos encontrar constantes de proporcionalidad para obtener una resultante dada, digamos $\mathbf{r} = (-1, -1)$. En otras palabras, supongamos que queremos encontrar números no negativos c_1, c_2, c_3 tales que

$$c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + c_3\mathbf{f}_3 = (-1, -1).$$

(Un valor negativo de alguno de los c invertiría la dirección de la fuerza correspondiente.) Esta ecuación vectorial es equivalente al sistema de ecuaciones que se obtiene sustituyendo los vectores dados (2) que determinan las direcciones de las fuerzas. Tenemos que

$$c_1\begin{pmatrix}-1\\3\end{pmatrix} + c_2\begin{pmatrix}4\\3\end{pmatrix} + c_3\begin{pmatrix}-2\\-4\end{pmatrix} = \begin{pmatrix}-1\\-1\end{pmatrix} \tag{3}$$

o bien

$$-c_1 + 4c_2 - 2c_3 = -1,$$
$$3c_1 + 3c_2 - 4c_3 = -1.$$

Como tenemos dos ecuaciones con tres incógnitas, es de esperar que, en general, si especificamos uno de los c, podremos encontrar los otros. Sin embargo, recordemos que los c deben ser no negativos. (En particular, una mirada a la Fig. 2(a) muestra que no podríamos obtener una resultante igual a $(-1, 1)$, a no ser que c_3 fuera positivo.) Por tanto, tratamos con $c_3 = 1$. Esto conduce a las ecuaciones

$$-c_1 + 4c_2 = 1$$
$$c_1 + c_2 = 1,$$

que tienen la solución $c_1 = \frac{3}{5}$, $c_2 = \frac{2}{5}$. Entonces el triple $(c_1, c_2, c_3) = (\frac{3}{5}, \frac{2}{5}, 1)$ es una solución posible y los tres vectores fuerza son

$$c_1\mathbf{f}_1 = (-\tfrac{3}{5}, \tfrac{9}{5}), \qquad c_2\mathbf{f}_2 = (\tfrac{8}{5}, \tfrac{6}{5}), \qquad c_3\mathbf{f}_3 = (-2, -4),$$

cuyas magnitudes son

$$|c_1\mathbf{f}_1| = \tfrac{3}{5}\sqrt{10}, \qquad |c_2\mathbf{f}_2| = 2, \qquad |c_3\mathbf{f}_3| = 2\sqrt{5}.$$

En forma análoga, podríamos buscar una asignación de fuerzas que pusiera en **equilibrio** el sistema, es decir, tal que la resultante sea el vector cero. Entonces habríamos remplazado el vector $(-1, -1)$ en el lado derecho de la Ec. (2) por $(0, 0)$, resolviendo el nuevo sistema de la misma manera.

Ejemplo 3. En este ejemplo consideramos la idea de un **camino aleatorio**, en el cual un objeto se mueve entre un número finito de posiciones a lo largo de caminos especificados, teniendo cada camino una cierta probabilidad de ser usado. En forma específica, suponemos que la probabilidad de abandonar una cierta posición a lo largo de un camino determinado es la misma para todos los caminos que salen de esa posición. Algunas configuraciones posibles se muestran en la Fig. 3 y es claro que ellas son redes.

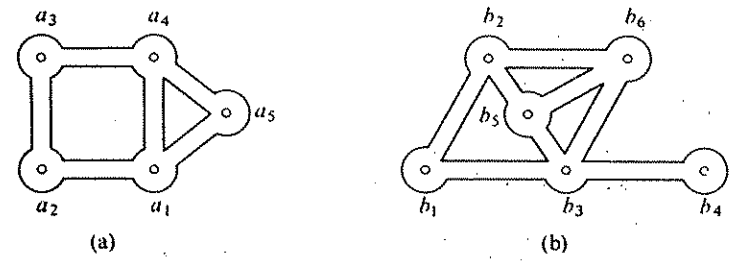


Figura 3

Se subentiende que una probabilidad es un número entre 0 y 1 y que la probabilidad de un evento particular es igual a la suma de las probabilidades de las distintas maneras en que puede suceder este evento. Así, en la Fig. 3(a), como hemos supuesto que todos los caminos que salen de a_5 son igualmente probables, sucede que la probabilidad de salir de a_5 por uno de los dos caminos posibles es $\frac{1}{2}$. De la misma forma, cada uno de los caminos que salen de a_4 tiene la probabilidad $\frac{1}{3}$. Suponemos, además, que el paso de una posición a otra es tal, que la probabilidad de dos eventos sucesivos es igual al producto de sus probabilidades respectivas. Así, la probabilidad de ir de a_2 a a_1 y luego a a_4 es $(\frac{1}{2})(\frac{1}{3}) = \frac{1}{6}$. Ahora nos hacemos preguntas como la siguiente: ¿Cuál es la probabilidad p_k de que partiendo de a_k se llegue a a_5 *sin pasar* por a_4? Vemos que si partimos de a_1 podemos ir directamente de a_1 a a_5 con probabilidad $\frac{1}{3}$ o podemos ir a a_2 con probabilidad $\frac{1}{3}$ y *luego* ir a a_5 con probabilidad p_2. Por tanto,

$$p_1 = \tfrac{1}{3} + (\tfrac{1}{3})p_2$$

Análogamente, ya que no se puede pasar por a_4,

$$p_2 = (\tfrac{1}{2})p_1 + (\tfrac{1}{2})p_3$$

$$p_3 = (\tfrac{1}{2})p_2.$$

Podemos reescribir estas ecuaciones en la forma

$$p_1 - (\tfrac{1}{3})p_2 = \tfrac{1}{3}$$

$$-(\tfrac{1}{2})p_1 + p_2 - (\tfrac{1}{2})p_3 = 0$$

$$-(\tfrac{1}{2})p_2 + p_3 = 0$$

y resolverlas de la manera habitual. Obtenemos

$$p_1 = \tfrac{3}{7}, \qquad p_2 = \tfrac{2}{7}, \qquad p_3 = \tfrac{1}{7}.$$

Sucede que mientras más cerca de a_5 empecemos es más probable llegar a a_5 sin pasar por a_4. Sin embargo, es importante comprender que, en general, las probabilidades dependen de la configuración global.

Un análisis como el que hemos hecho sería útil para distinguir entre el comportamiento totalmente aleatorio de una criatura en un laberinto y un comportamiento inteligente o condicionado.

Ejemplo 4. Un sistema interconectado de cañerías de agua puede ser considerado como una red en la cual los nodos son las uniones. Es habitual asignar una dirección de flujo a cada una de las cañerías en este tipo de redes, indicada por flechas en la Fig. 4. Entonces, un número positivo r indicará una velocidad de flujo en la dirección asignada, mientras que un número negativo $-r$ representará un flujo con la misma velocidad, pero en la dirección opuesta. Separaremos las velocidades de flujo en velocidades internas r_k y externas que denotaremos por t_k o s_k. (Asignar una velocidad externa 0 a una juntura es lo mismo que cerrar la cañería externa en esa unión.) Suponemos que en cada una de las junturas, el flujo de entrada es igual al flujo de salida. Así, por ejemplo, en

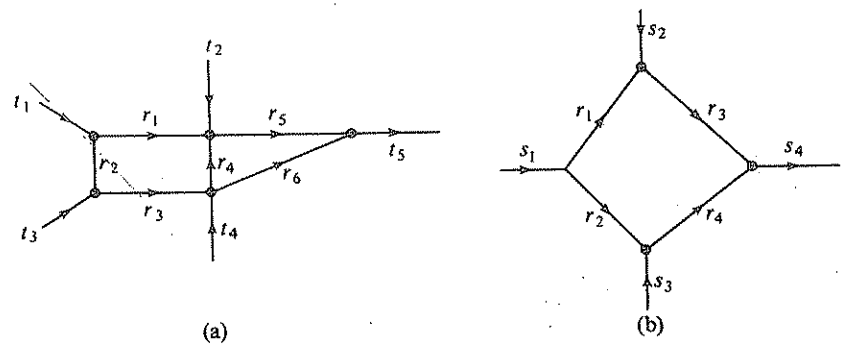


Figura 4

la esquina superior izquierda de la Fig. 4(a) tenemos que $t_1 = r_1 + r_2$, mientras que en la esquina inferior izquierda se debe tener $t_3 + r_2 = r_3$. En el caso de la red de la Fig. 4(a), el conjunto completo de ecuaciones que relacionan las r y las t puede ser expresado en la forma

$$\begin{aligned} r_1 + r_2 &= t_1 \\ -r_1 \quad -r_4 + r_5 &= t_2 \\ -r_2 + r_3 &= t_3 \\ -r_3 + r_4 \quad + r_6 &= t_4 \\ r_5 + r_6 &= t_5 \end{aligned} \tag{4}$$

En forma vectorial podemos escribir $A\mathbf{r} = \mathbf{t}$, donde A es la matriz 5-por-6

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

La ecuación vectorial muestra que hay una relación lineal entre $\mathbf{t}$ y $\mathbf{r}$, a saber, $\mathbf{t}$ es una función lineal de $\mathbf{r}$. Esto implica que si los flujos r_k son multiplicados por un número c, entonces también los flujos t_k son multiplicados por c. Igualmente, si se superimponen dos conjuntos de flujos r_k y r'_k, entonces los flujos externos resultantes están dados por $\mathbf{t} = A(\mathbf{r} + \mathbf{r}')$. Este fenómeno se llama el **principio de superposición** para sistemas lineales. Evidentemente, si especificamos los $\mathbf{r}$, los $\mathbf{t}$ quedarán totalmente determinados.

Dando vuelta el problema, podemos preguntarnos hasta qué punto la especificación de los flujos externos en las junturas determinará los flujos r_k en las cañerías. En particular, podemos tratar de exigir que cada uno de los flujos

externos t_k sea cero. Esto conduce al sistema de cinco ecuaciones con seis incógnitas

$$\begin{aligned} r_1 + r_2 &= 0 \\ -r_1 \qquad -r_4 + r_5 &= 0 \\ -r_2 + r_3 &= 0 \\ -r_3 + r_4 \qquad r_6 &= 0 \\ r_5 + r_6 &= 0. \end{aligned}$$

Podemos tomar $r_6 = a$, un número cualquiera, de donde obtenemos que $r_5 = -a$. Ahora hacemos $r_4 = b$, un número cualquiera. Las cuatro ecuaciones restantes pueden ser resueltas fácilmente en términos de a y b, obteniéndose

$$\mathbf{r}_0 = (-a - b, a + b, a + b, b, -a, a).$$

Una verificación sencilla es que estos valores de r_1 a r_6 satisfacen el sistema anterior. Por tanto, hay un número infinito de flujos distintos que producen flujos externos nulos en cada juntura. En forma análoga, podemos buscar una solución $\mathbf{r}$ del sistema original (4) para un vector $\mathbf{t}$ dado. Entonces, por el principio de superposición, toda solución $\mathbf{r}_0$ puede ser sumada a $\mathbf{r}$ obteniéndose una nueva solución $\mathbf{r}_0 + \mathbf{r}$. En efecto,

$$\begin{aligned} A(\mathbf{r}_0 + \mathbf{r}) &= A\mathbf{r}_0 + A\mathbf{r} \\ &= 0 + \mathbf{t} = \mathbf{t}. \end{aligned}$$

Ejemplo 5. La demostración de la regla de Simpson para integrar aproximadamente está basada en el requisito de que debe dar resultados exactos al aplicarla a polinomios cuadráticos. La regla de Simpson da una aproximación para la integral de una función en el intervalo $(a - h, a + h)$ en términos del valor de la función en los puntos externos y medio del intervalo en la forma

$$\int_{a-h}^{a+h} f(x)\,dx \approx uf(a - h) + vf(a) + wf(a + h)$$

donde u, v y w son constantes. Si se quiere que la fórmula sea correcta para polinomios de grado menor o igual que 2, en particular debe serlo para los polinomios $f_0(x) = 1$, $f_1(x) = x$ y $f_2(x) = x^2$. Cada uno de éstos da una ecuación para u, v y w. Por ejemplo, para $f_0(x) = 1$, tenemos que

$$\int_{a-h}^{a+h} f_0(x)\,dx = 2h \qquad \text{y} \qquad f_0(a - h) = f_0(a) = f_0(a + h) = 1$$

de módo que $u + v + w = 2h$. Usando $f_1(x) = x$ y $f_2(x) = x^2$ se obtiene

$$(a - h)u + av + (a + h)w = 2ah$$

y

$$(a^2 + h^2 - 2ah)u + a^2v + (a^2 + h^2 + 2ah)w = 2a^2h + \tfrac{2}{3}h^3.$$

Es fácil resolver este sistema (véase Ejercicio 13) y el resultado es $u = w = \frac{1}{3}h$, $v = \frac{4}{3}h$.

Hemos obtenido una regla que es correcta para los polinomios f_0, f_1 y f_2. Su exactitud para polinomios cuadráticos cualesquiera es clara. Sea $E(f)$ el error cuando se aplica esta regla a una función f, de modo que

$$E(f) = \int_{a-h}^{a+h} f(x)\,dx - \tfrac{1}{3}hf(a - h) - \tfrac{4}{3}f(a) - \tfrac{1}{3}f(a + h).$$

Es fácil comprobar que E es una función lineal del espacio vectorial de los polinomios con valores en $\mathfrak{R}^1$ y $E(f_0) = E(f_1) = E(f_2) = 0$ por construcción. Si f es un polinomio cuadrático $f(x) = px^2 + qx + r$, entonces $f = pf_2 + qf_1 + rf_0$ y por linealidad,

$$E(f) = pE(f_2) + qE(f_1) + rE(f_0) = 0.$$

El mismo método puede ser aplicado para obtener una gran cantidad de fórmulas útiles en el análisis numérico (véase Ejercicio 14).

EJERCICIOS

1. La Fig. 1(b) muestra una red eléctrica con sus resistencias (en ohmios) marcadas en cada línea. Suponga que una fuente externa de poder mantiene el nodo A con un potencial de 10 voltios y el nodo B con un potencial de 4 voltios. Use el método del Ejemplo 1 para encontrar ecuaciones para el potencial en cada uno de los otros nodos y resuelva las ecuaciones. Usando los resultados calcule la corriente que fluye a la red en A.

2. Los lados y vértices de un cubo tridimensional forman una red con 8 nodos y 12 lados. Suponga que cada lado es un alambre con 1 ohm de resistencia y que dos de los vértices tienen conexiones externas que mantienen uno de ellos con un potencial 1 y el otro con un potencial 0. Encuentre el valor del potencial en los otros vértices y la corriente que fluye en las conexiones externas si
   (a) Los dos vértices con conexiones externas están en esquinas opuestas del cubo.
   (b) Si ellos están en los extremos de un lado.
   (c) Están en esquinas opuestas de una de las caras del cubo.

3. (a) Suponga que tres fuerzas que actúan en el origen de $\mathfrak{R}^3$ tienen la misma dirección que (1, 0, 0), (1, 1, 0) y (1, 1, 1). Encuentre las magnitudes de las fuerzas que, actuando en esas direcciones, tienen un vector fuerza resultante igual a $(-1, 2, 4)$.
   (b) ¿Es posible que *cualquier* vector fuerza sea resultante de fuerzas que actúan en las direcciones indicadas en la parte (a)?

4. Demuestre que si se usan tres vectores linealmente independientes en $\mathfrak{R}^3$ para especificar las direcciones de tres fuerzas, entonces siempre es posible encontrar magnitudes para estas fuerzas, de modo que la resultante de esas fuerzas sea cualquier vector dado.

5. Si tres fuerzas en $\mathfrak{R}^2$ actúan en las direcciones de (2, 1), (2, 2) y $(-3, -1)$ demuestre que se pueden encontrar magnitudes de modo que el sistema esté en equilibrio.

6. Suponga que se tiene un camino aleatorio en la red indicada en la Fig. 3(a). ¿Cuál es la probabilidad p_k, $k = 1, 2, 3$, que el camino empiece en a_k y vaya a a_4 sin pasar por a_5?

7. (a) Suponga que una partícula hace un camino aleatorio en la red de la Fig. 3(b). Si p_k es la probabilidad de ir de b_k a b_6 sin pasar por b_5, encuentre p_k para $k = 1, 2, 3, 4$.
   (b) ¿Cómo se modifica el resultado de la parte (a) si se elimina b_4 y el segmento que lo une a b_3?
   (c) ¿Cómo se modifica el resultado de (a) si se agrega un nuevo segmento entre b_4 y b_6?

8. Suponga que se pueden hacer varias mezclas con sustancias S_1, S_2, S_3, S_4 que tienen densidades 2, 17, 3 y 1 respectivamente, medidas en gramos por centímetro cúbico. Suponga también que el precio en centavos por gramo de estas sustancias es 2, 3, 1 y 5 respectivamente. ¿Es posible hacer una mezcla que pese 10 g., con un volumen de 20 cm^3 y que cueste 2 dólares? ¿Qué sucede si se termina S_2?

9. Suponiendo que la función lineal f de $\mathcal{R}^3$ en $\mathcal{R}^3$ es tal que $f(\mathbf{e}_1) = (1, 2, -1)$, $f(\mathbf{e}_2) = (0, 1, 1)$ y $f(\mathbf{e}_3) = (1, 2, 1)$. Encuentre vectores $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ tales que $f(\mathbf{x}_k) = \mathbf{e}_k$ para $k = 1, 2, 3$.

10. Sea f la función lineal de $\mathcal{R}^3$ en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix}.$$

Demuestre que los puntos $\mathbf{x}$ de $\mathcal{R}^3$ tales que $f(\mathbf{x}) = 0$ están en una recta.

11. (a) Suponga que el vector $\mathbf{t} = (t_1, t_2, t_3, t_4, t_5)$ de la Fig. 4(a) es $t = (-1, 0, 1, 2, 1)$. Encuentre un vector $\mathbf{r}$ que determine velocidades de flujo internas compatibles.
    (b) Verifique que el vector $\mathbf{r}_0 = (-a - b, a + b, a + b, b, -a, 0)$ para todo a y b es compatible con el flujo externo $\mathbf{t} = 0$ en la Fig. 4(a).

12. (a) Suponga que el vector de flujo externo de la Fig. 4(b) está dado por $\mathbf{t} = (1, 1, 2, 4)$. Demuestre que hay más de un vector flujo interno $\mathbf{r}$ compatible y encuentre dos de ellos.
    (b) Suponga que el vector de flujo externo de la Fig. 4(b) está dado por $\mathbf{t} = (1, 0, 1, 1)$. Demuestre que no hay un vector flujo interno compatible.

13. Encuentre las soluciones u, v, w del Ejemplo 5. [*Sugerencia*: Empiece por restar a veces la primera ecuación de la segunda y a^2 veces la primera a la tercera.]

14. Use el método del Ejemplo 5 para encontrar constantes t, u, v, w tales que la fórmula

$$\int_a^{a+3h} f(x)\,dx = tf(a) + uf(a + h) + vf(a + 2h) + wf(a + 3h)$$

es exacta si f es un polinomio de grado menor o igual que 3.

3. Teoría de las ecuaciones lineales

En esta sección nos preocuparemos de la formulación del método para resolver sistemas de ecuaciones lineales que ha sido ilustrado en las secciones precedentes y de demostrar que él funciona siempre. Empezamos por dar una definición precisa del principio de la «no interferencia», discutida informalmente después

del Ejemplo 3 de la Sec. 1. En ese lugar expresamos esta idea diciendo que cada ecuación de un sistema contenía una variable que no aparecía en otra ecuación. Las definiciones que siguen expresan esta propiedad de «no interferencia» en términos de la matriz de coeficientes.

Un elemento de una matriz se llama **elemento principal** si él es el primer elemento no cero de una fila. Así, por ejemplo, en

$$\begin{pmatrix} 3 & 0 & 2 & 4 \\ 0 & 0 & 4 & -10 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 3 & 1 \end{pmatrix},$$

los elementos principales son el primero en la primera fila, el tercero en la segunda fila y el primero en la cuarta fila. No hay elemento principal en la tercera fila, ya que todos ellos son cero. Decimos que una matriz es **reducida** si:

3.1 (a) Toda columna que contiene un elemento principal es cero, excepto por el elemento principal.
    (b) Todos los elementos principales son 1.

Al discutir las matrices reducidas es a menudo necesario distinguir entre las columnas que contienen elementos principales y las que no contienen ninguno. (Por supuesto, una fila contiene un elemento principal si, y solo si, ella es distinta de cero.) Diremos que las columnas que contienen elementos principales son **pivotales.** En una matriz reducida, todo elemento principal pertenece a una fila distinta de cero y a una columna pivotal; diremos que la fila y la columna son **asociadas.** Esto establece una relación biunívoca entre filas distintas de cero y columnas pivotales y prueba el siguiente hecho al cual nos referiremos más adelante.

3.2 En una matriz reducida, el número de columnas pivotales es igual al número de filas distintas de cero.

Ejemplo 1. Consideremos las matrices

$$A = \begin{pmatrix} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 3 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A es reducida, sus columnas pivotales son la primera (asociada a la primera

fila) y la tercera (asociada a la segunda fila). B también es reducida, sus columnas pivotales son la primera, la tercera y la cuarta y ellas son asociadas a la segunda, primera y tercera filas, respectivamente. La matriz C no cumple la condición (b), pero multiplicando su primera fila por $\frac{1}{2}$ se obtiene una matriz reducida. D no satisface la condición (a), pero si restamos 5 veces la tercera fila de la primera y después sumamos 2 veces la tercera fila a la segunda, obtenemos una matriz reducida.

Damos, a continuación, algunos teoremas acerca de la resolución de sistemas de ecuaciones cuya matriz de coeficientes es reducida. El primer teorema muestra cómo determinar si una ecuación tiene soluciones.

3.3 Teorema

Sea A una matriz reducida. Si A tiene una fila de ceros tal que el coeficiente correspondiente de $\mathbf{b}$ no es cero, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ no tiene solución. En caso contrario, la ecuación tiene por lo menos una solución.

Demostración. Si una fila de A es cero, entonces el coeficiente correspondiente de $A\mathbf{x}$ es cero para todo $\mathbf{x}$. Por tanto, si el coeficiente correspondiente de $\mathbf{b}$ no es cero, no puede haber una solución.

Para demostrar la segunda parte del teorema, supongamos que a cada fila cero de A (si hay alguna) corresponde un coeficiente cero en $\mathbf{b}$, mostraremos cómo obtener un vector $\mathbf{x}$ que es una solución. El vector $\mathbf{x}$ debe tener coordenadas $x_1, \ldots, x_n$ que corresponden a las n columnas de A, mientras que $\mathbf{b}$ tiene coeficientes $b_1, \ldots, b_m$ que corresponden a las m filas de A. Si la columna j de A no es pivotal, pongamos $x_j = 0$. Si la columna j de A es pivotal, sea i su fila asociada, entonces ponemos $x_j = b_i$. Afirmamos que el vector $\mathbf{x}$ así construido satisface $A\mathbf{x} = \mathbf{b}$. El coeficiente i del producto es $\sum_{k=1}^{n} a_{ik}x_k$. Debemos demostrar que es igual a b_i. Este número es cero si la fila i de A es cero, pero en este caso b_i también es cero. Si la fila i no es cero, entonces $a_{ij} = 1$, donde j es la columna pivotal asociada. Por construcción, $x_j = b_j$, de modo que $a_{ij}x_j = b_j$. Los términos $a_{ik}x_k$ con $k \neq j$ son todos cero, ya que $a_{ik} = 0$ si la columna k es pivotal (puesto que A es reducida) y $x_k = 0$ si la columna k no es pivotal (por construcción).

Ejemplo 2. Consideremos

$$A = \begin{pmatrix} 0 & 0 & 1 & 5 & 0 \\ 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{b}_1 = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 3 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 3 \end{pmatrix}.$$

La tercera fila de A es cero y el tercer coeficiente de $\mathbf{b}_1$ no lo es, por tanto, de acuerdo con el Teorema 3.3, la ecuación $A\mathbf{x} = \mathbf{b}_1$ no tiene solución. Por otra parte, el tercer coeficiente de $\mathbf{b}_2$ es cero.

Las columnas pivotales de A son la primera, la tercera y la quinta y sus filas asociadas son, respectivamente, la segunda, la primera y la cuarta. La demostración del teorema muestra que si tomamos

$$\mathbf{x} = \begin{pmatrix} 4 \\ 0 \\ 1 \\ 0 \\ 3 \end{pmatrix}$$

tal que su primer, tercer y quinto coeficiente es igual al segundo, primero y cuarto coeficiente de $\mathbf{b}_2$, respectivamente, y los otros coeficientes son cero, entonces $A\mathbf{x}$ será igual a $\mathbf{b}_2$. Esto se comprueba fácilmente efectuando la multiplicación matricial.

El problema siguiente es determinar si una ecuación que tiene una solución tiene más de una. El Teorema 3.4 y su corolario muestran que este problema puede ser reducido a un caso especial y luego el Teorema 3.5 trata sobre este caso especial.

3.4 Teorema

Supongamos que $\mathbf{x}_0$ es una solución de la ecuación $A\mathbf{x} = \mathbf{b}$. Entonces, $\mathbf{x}_1$ es también una solución si, y solo si, $\mathbf{x}_1 - \mathbf{x}_0$ es una solución de $A\mathbf{x} = 0$.

Demostración. Tenemos que $A(\mathbf{x}_1 - \mathbf{x}_0) = A\mathbf{x}_1 - A\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = 0$. Conversamente, si $A(\mathbf{x}_1 - \mathbf{x}_0) = 0$, entonces $A\mathbf{x}_1 = A\mathbf{x}_0 = \mathbf{b}$.

La ecuación $A\mathbf{x} = 0$ se llama a menudo la ecuación **homogénea** asociada a la ecuación $A\mathbf{x} = \mathbf{b}$. Nótese que la ecuación homogénea tiene siempre por lo menos una solución, a saber $\mathbf{x} = 0$. Usando 3.4, obtenemos de inmediato el siguiente

Corolario

Si $A\mathbf{x} = \mathbf{b}$ tiene *por lo menos* una solución, entonces tiene una solución si, y solamente si, $\mathbf{x} = 0$ es la única solución de la ecuación homogénea asociada.

3.5 Teorema

Sea A una matriz reducida. Si toda columna de A es pivotal, entonces $\mathbf{x} = 0$ es la única solución de la ecuación homogénea $A\mathbf{x} = 0$. En caso contrario, la ecuación tiene soluciones $\mathbf{x} \neq 0$.

Demostración. Supongamos que todas las columnas de A son pivotales. Sea la fila i asociada a la columna j. Entonces $a_{ij} = 1$ y $a_{ik} = 0$ para $k \neq j$ (pues-

to que A es reducida y todas sus columnas son pivotales). Por tanto, el coeficiente i de $A\mathbf{x}$ es $\sum_{k=1}^{n} a_{ik}x_k = x_j$; de modo que si $A\mathbf{x} = 0$, entonces $x_j = 0$ para todo j.

Conversamente, la columna r no es pivotal, podemos construir una solución distinta de cero de $A\mathbf{x} = 0$ como sigue. Sea $x_r = 1$ (lo que garantiza que $\mathbf{x} \neq 0$) y $x_j = 0$ si la columna j es otra columna no pivotal. Si la columna j es pivotal, sea $x_j = -a_{ir}$, si la fila i es asociada a la columna j. Igual que en la demostración del Teorema 3.4, miramos al producto $A\mathbf{x}$ fila por fila. Las filas cero dan cero en el producto. Si la fila i de A no es cero, obtenemos la suma $\sum_{k=1}^{n} a_{ik}x_k$. Si la columna k es pivotal y $k \neq j$ (la columna j es asociada a la fila i), tenemos que $a_{ik} = 0$. Si ella no es pivotal y $k \neq r$, tenemos que $x_k = 0$. Por tanto, esta suma queda reducida a $a_{ij}x_j + a_{ir}x_r = 1 \cdot (-a_{ir}) + a_{ir} \cdot 1 = 0$, como se deseaba.

Ejemplo 3. Consideremos la ecuación homogénea $A\mathbf{x} = 0$, donde A es la matriz del Ejemplo 2. La segunda y la cuarta columna de A no son pivotales. La construcción usada en la demostración del Teorema 3.5 puede ser usada para obtener una solución con $x_2 = 1$, $x_4 = 0$ y otra con $x_4 = 1$, $x_2 = 0$. Los vectores obtenidos son

$$\mathbf{y} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{y} \quad \mathbf{z} = \begin{pmatrix} -3 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix}.$$

Cualquier combinación $r\mathbf{y} + s\mathbf{z}$ es también una solución. (¿Por qué?) Usando el Teorema 3.4 para combinar esta información con la del Ejemplo 2, vemos que todos los vectores de la forma

$$\begin{pmatrix} 4 \\ 0 \\ 1 \\ 0 \\ 3 \end{pmatrix} + r \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -3 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix}$$

son soluciones de

$$\begin{pmatrix} 0 & 0 & 1 & 5 & 0 \\ 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 3 \end{pmatrix}$$

Hemos resuelto la mayor parte de las preguntas acerca de las soluciones de sistemas lineales cuya matriz de coeficientes es reducida. Si podemos transformar cualquier sistema en uno reducido, usando operaciones elementales, entonces tendremos un método general para encontrar las soluciones de un sistema lineal cualquiera. Los tres primeros ejemplos de esta sección ilustran un procedimiento que puede ser aplicado con éxito a cualquier matriz.

Supongamos que una matriz no es reducida. Entonces debe existir una columna cuyo elemento principal es tal que no cumple 3.1(a) o 3.1(b) (o ambas). Si la columna contiene el elemento principal r para la fila i, multiplicando la fila i por r^{-1} se obtiene un elemento principal 1. (Como r era un elemento principal, no puede ser cero. Por supuesto, podría ser 1, en cuyo caso la multiplicación sería innecesaria.) Si otros elementos en la columna no son ceros, ellos se pueden hacer cero sumando múltiplos adecuados de la fila i a la fila en que ellos están. Si una columna era «correcta» antes de efectuar estas operaciones, ella deberá tener un cero en su elemento 1 y, por tanto, no se altera por las operaciones efectuadas. Así hemos descrito un procedimiento que puede ser aplicado a una matriz no reducida, que aumenta el número de columnas que satisfacen las condiciones 3.1. Si la matriz resultante no es reducida, se puede repetir el proceso. Después de a lo más n pasos, se obtendrá una matriz reducida, donde n es el número de columnas de la matriz. Hemos demostrado

3.6 Teorema

Dada una matriz A, se puede encontrar una sucesión de operaciones elementales que transforman a A en una matriz reducida.

Ejemplo 4. Aplicaremos el método de la demostración del Teorema 3.6 para reducir la matriz

$$A = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 2 & 6 & 0 & -1 \\ -1 & -3 & 4 & 1 \end{pmatrix}.$$

La columna 1 no satisface 3.1(a), pero tiene elemento principal 1 en la primera fila, de modo que no es necesario efectuar una multiplicación elemental. Restando 2 veces la fila 1 a la fila 2 y sumando la fila 1 a la fila 3 se hacen cero los otros dos elementos y se tiene

$$A_2 = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 0 & 4 & -1 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

La columna 2 no tiene elemento principal. El 4 en la columna 3 puede ser convertido en un 1 multiplicando la segunda fila por $\frac{1}{4}$, con lo cual se obtiene

$$A_3 = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 0 & 1 & -\frac{1}{4} \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

Sumando 2 veces la fila 2 a la fila 1 y (-2) veces la fila 2 a la fila 3 se obtienen ceros en la columna 3:

$$A_4 = \begin{pmatrix} 1 & 3 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{4} \\ 0 & 0 & 3 & \frac{3}{2} \end{pmatrix}.$$

Multiplicando por $\frac{2}{3}$ la fila 3 y luego sumando $\frac{1}{2}$ veces la fila 3 a la fila 1 y $\frac{1}{4}$ veces la fila 3 a la fila 2 se obtiene la matriz reducida

$$\begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

como resultado final.

Al resolver este ejemplo, hemos seguido el procedimiento estándar dado en la demostración de 3.6. Por supuesto, este camino no es el único que conduce a una matriz reducida, y una sucesión diferente puede requerir menos cálculos aritméticos. Por ejemplo, consideremos la matriz A_2. Sumando la fila 3 a la fila 1 y restando 2 veces la fila 3 de la fila 2 se tiene

$$\begin{pmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

Luego sumando $(\frac{1}{3})$ veces la fila 2 a la fila 1 y $(\frac{1}{3})$ veces la fila 2 a la fila 3 se obtiene

$$\begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 2 & 0 \end{pmatrix}.$$

Finalmente, multiplicando la fila 2 por $(-\frac{1}{3})$ y la fila 3 por $(\frac{1}{2})$ se obtiene la matriz reducida

$$\begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Veremos ahora el caso especial de n ecuaciones con n incógnitas, es decir, de sistemas con matrices cuadradas de coeficientes. El siguiente teorema es el más importante para las aplicaciones.

3.7 Teorema

Si A es una matriz cuadrada tal que la ecuación homogénea $A\mathbf{x} = 0$ tiene la única solución $\mathbf{x} = 0$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene exactamente una solución para cada vector $\mathbf{b}$.

Demostración. Por el Teorema 3.6, podemos convertir A en una matriz reducida C y, usando las mismas operaciones lineales, convertir $\mathbf{b}$ en el vector $\mathbf{d}$ de modo que $C\mathbf{x} = \mathbf{d}$ tiene el mismo conjunto de soluciones que $A\mathbf{x} = \mathbf{b}$. Por tanto, la hipótesis implica que $C\mathbf{x} = 0$ tiene la única solución $\mathbf{x} = 0$. Como C es reducida, el Teorema 3.5 implica que todas las columnas de C son pivotales. El número de filas distintas de cero de C es igual al número de columnas pivotales. Como C es cuadrada, no puede haber filas cero y, por el Teorema 3.3, $C\mathbf{x} = \mathbf{d}$ (y, por tanto, $A\mathbf{x} = \mathbf{b}$) tiene por lo menos una solución.

Del corolario del Teorema 3.3 se deduce que hay exactamente una solución. Ahora podemos demostrar el converso del Teorema 1.3, es decir.

3.8 Teorema

Si A es invertible, entonces ella puede ser transformada en I mediante operaciones elementales.

Demostración. Si existe $\mathbf{x} \neq 0$ tal que $A\mathbf{x} = 0$, A no puede ser invertible, ya que $A\mathbf{x} = 0$ implicaría $A^{-1}A\mathbf{x} = A^{-1}0$, es decir, $\mathbf{x} = 0$, lo cual sería una contradicción. En caso contrario, A puede ser transformada en una matriz reducida C con todas las columnas pivotales. Cada columna de C tiene un 1 y todos los demás elementos son 0 y los 1 en las distintas columnas están en filas distintas. Cambiando en forma adecuada el orden de las filas (lo cual puede ser hecho mediante transformaciones elementales) obtendremos la matriz identidad. Por tanto, A puede ser transformada en I mediante operaciones elementales.

Vemos así que es razonable aplicar el método usado para calcular inversos dado por el Teorema 1.3 a cualquier matriz cuadrada A. Si el método tiene éxito, se ha demostrado que A es invertible y se ha calculado A^{-1}. Si el método falla, se obtiene una matriz reducida con a lo menos una columna no pivotal y, por tanto, usando el Teorema 3.5, un $\mathbf{x}$ distinto de cero tal que $A\mathbf{x} = 0$, lo cual demuestra que A no es invertible.

Concluiremos notando que toda operación elemental, al ser aplicada a un vector columna en $\mathcal{R}^n$, es una función lineal, llamada una **transformación elemental**. En efecto, es obvio que: (a) la multiplicación de una coordenada por $r \neq 0$, (b) la suma de un múltiplo de una coordenada a otra y (c) la transposición de dos coordenadas, son operaciones lineales. Por tanto, cada una de estas operaciones puede ser efectuada mediante la multiplicación por la izquierda por una **matriz elemental** n-por-n. (La forma precisa de estas matrices se describe en el Ejercicio 7.) Esto nos permite demostrar el siguiente teorema.

3.9 Teorema

Una matriz invertible A puede ser escrita como un producto de matrices elementales. Por tanto, la función lineal definida por

$$f(\mathbf{x}) = A\mathbf{x}$$

puede ser expresada como una composición de transformaciones elementales.

Demostración. Por el Teorema 3.8, la matriz A puede ser transformada en I aplicando a A un producto

$$Q = A_1 \dots A_k$$

de matrices elementales. Entonces, $QA = I$. Pero como cada operación elemental es reversible, cada matriz A_i tiene un inverso A_i^{-1}. Entonces Q tiene inverso

$$Q^{-1} = A_k^{-1} \dots A_1^{-1},$$

y $A = Q^{-1}I = A_k^{-1} \dots A_1^{-1}$, lo cual expresa A en la forma deseada.

EJERCICIOS

1. Determine cuáles de las matrices siguientes son reducidas. Para aquellas que no lo son, explique exactamente cuál de las condiciones no se cumple. Para las que son reducidas, indique cuáles son sus columnas pivotales y cuáles son sus filas asociadas.

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Para cada una de las matrices del Ejercicio 1 que no sea reducida encuentre una operación elemental que la haga reducida.

3. Sean

$$\mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \qquad \mathbf{s} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{t} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Para cada una de las ecuaciones $A\mathbf{x} = \mathbf{r}$, $A\mathbf{x} = \mathbf{s}$, $A\mathbf{x} = \mathbf{t}$, donde A es la matriz del Ejercicio 1, determine si la ecuación tiene una, ninguna o más de una solución. Si hay una solución, encuéntrela, si hay más de una, encuentre dos.

4. Repita el Ejercicio 3 usando las matrices B, C y D del Ejercicio 1 en lugar de A. (Recuerde que debe usar la matriz de coeficientes en forma reducida.)

5. Demuestre que si una matriz n-por-n es reducida y no tiene ninguna fila cero, entonces cada fila y cada columna tiene $n - 1$ ceros y un uno. Demuestre, por tanto, que puede ser convertida en la matriz identidad usando transposiciones elementales.

6. Determine los conjuntos de soluciones de los siguientes sistemas de ecuaciones.

(a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 10 \\ 11 \\ 12 \end{pmatrix}.$ [*Resp*: $r(1, -2, 1) + (0, -9, \frac{28}{3})$.]

(b) $x_1 + x_2 = 1$
$x_2 + x_3 = 2$
$x_3 + x_4 = 3$
$x_4 + x_1 = 4.$

(c) $\begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -4 \\ 1 & 3 & 9 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 & 2 \\ 6 & 2 \\ 0 & 0 \end{pmatrix}.$

(d) $x_1 + 2x_2 + x_3 - x_4 = 1$
$x_2 - x_3 + x_4 = 0$
$x_1 + 3x_3 - 3x_4 = 3.$

(e) $x_1 + 2x_2 + x_3 - x_4 = 1$
$x_2 - x_3 + x_4 = 0$
$x_1 + 3x_3 - 3x_4 = 1.$

7. (a) Denote por $D_i(r)$ la matriz que es igual a la identidad salvo que su elemento i-ésimo en la diagonal es r en lugar de 1. Demuestre que la matriz producto

$$D_i(r)M$$

es la modificación elemental de M que se obtiene al multiplicar la i-ésima fila de M por r.

(b) Denote por E_{ij} la matriz con elementos 0 salvo el elemento ij que es 1. Por ejemplo, para matrices 3-por-3.

$$E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Demuestre que si M es una matriz m-por-n y E_{ij} e I son m-por-m, entonces

$$(I + rE_{ij})M$$

es la modificación elemental consistente en sumar r veces la fila j a la fila i.

(c) Sea T_{ij} la matriz que se obtiene de I al intercambiar las filas i y j. Demuestre que si se intercambian las filas i y j de M se obtiene

$$T_{ij}M.$$

(d) Una matriz de la forma $(I + rE_{ij})$, $D_i(r)$ o T_{ij} se llama una **matriz elemental.** Demuestre que las matrices elementales son invertibles verificando que

$$(I + rE_{ij})^{-1} = (I - rE_{ij}), \qquad D_i(r)^{-1} = D_i\left(\frac{1}{r}\right), \quad \text{y} \quad T_{ij}^{-1} = T_{ij}.$$

8. Demuestre que si hay más incógnitas que ecuaciones en un sistema lineal, entonces hay infinitas soluciones o no hay ninguna.

9. Sea A una matriz cuadrada. Demuestre que A es invertible si, y solo si, toda ecuación $A\mathbf{x} = \mathbf{b}$ tiene *por lo menos* una solución.

10. Sea $p(x) = a_0 + a_1x + \ldots + a_nx^n$ un polinomio de grado $\leq n$. Un teorema conocido del álgebra afirma que si existen más de n valores de x tales que $p(x) = 0$, entonces todos los coeficientes son cero. Use este teorema para demostrar que si $x_0, \ldots, x_n$ son $n + 1$ números distintos y $b_0, \ldots, b_n$ son $n + 1$ números, entonces existe un solo polinomio de grado $\leq n$ tal que $p(x_0) = b_0, \ldots, p(x_n) = b_n$. [*Indicación*: Demuestre que el problema conduce a un sistema de ecuaciones lineales con incógnitas $a_0, \ldots, a_n$.]

11. Una matriz reducida en la cual la primera columna pivotal (partiendo de la izquierda) tiene como asociada a la primera fila, la segunda columna pivotal tiene como asociada a la segunda fila, etc., se dice que está en su forma **escalonada.** Demuestre que

(a) Toda matriz reducida puede ser puesta en forma escalonada mediante una sucesión de transposiciones elementales.

(b) En una matriz en forma escalonada, las filas cero (si hay alguna) son las últimas.

(c) Una matriz cuadrada en forma escalonada es la matriz identidad o tiene por lo menos una columna cero.

4. Espacios vectoriales, subespacios, dimensiones

En las primeras secciones de este capítulo nos hemos restringido a vectores en $\mathcal{R}^n$. En esta sección consideraremos espacios vectoriales más generales, a pesar de que algunas de las ideas ya han sido introducidas en el caso de $\mathcal{R}^2$ o $\mathcal{R}^3$. Nuestro fin es demostrar que todo espacio vectorial con una base con n elementos, tiene una dimensión bien definida.

Recordemos que un vector $\mathbf{x}$ es una **combinación lineal** de los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ si existen números $r_1, \ldots, r_n$ tales que

$$\mathbf{x} = r_1\mathbf{x}_1 + \ldots + r_n\mathbf{x}_n.$$

Ejemplo 1. Sean

$$\mathbf{x}_1 = (1, 0, 0), \quad \mathbf{x}_2 = (0, 1, 0), \quad \mathbf{x}_3 = (0, 0, 1), \quad \mathbf{x}_4 = (1, 1, 1).$$

Entonces $\mathbf{y} = (2, 2, 0)$ es una combinación lineal de $\mathbf{x}_1$ y $\mathbf{x}_2$, ya que $\mathbf{y} = 2\mathbf{x}_1 + 2\mathbf{x}_2$, también es una combinación lineal de $\mathbf{x}_3$ y $\mathbf{x}_4$, ya que $\mathbf{y} = 2\mathbf{x}_4 - 2\mathbf{x}_3$. Por otra parte, no es una combinación lineal de $\mathbf{x}_2$ y $\mathbf{x}_3$, ya que $r\mathbf{x}_2 + s\mathbf{x}_3$ tiene primera coordenada 0 para todo valor de r y s, de modo que es imposible que $\mathbf{y} = r\mathbf{x}_2 + s\mathbf{x}_3$.

Como $0 = 0\mathbf{x}_1 + \ldots + 0\mathbf{x}_n$, el vector cero es una combinación lineal de cualquier conjunto de vectores. Las combinaciones lineales de un solo vector $\mathbf{x}_1$ son los múltiplos numéricos $r\mathbf{x}_1$ de este vector.

Si un conjunto de vectores está en un plano en $\mathcal{R}^3$ que pasa por el origen, entonces toda combinación lineal de ellos está en el mismo plano. En efecto,

si $\mathbf{x}$ e $\mathbf{y}$ están en el plano que pasa por el origen, la regla del paralelogramo indica que la suma $\mathbf{x} + \mathbf{y}$ está en el mismo plano. Cualquier múltiplo numérico de $\mathbf{x}$ está en el mismo plano, ya que está en la recta que contiene a $\mathbf{x}$. Toda combinación lineal se forma efectuando multiplicaciones numéricas y sumas, con lo cual queda verificada la afirmación.

De igual manera, si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ están todos en una recta que pasa por el origen (de modo que todos ellos son múltiplos de un mismo vector), toda combinación lineal de ellos está en la misma recta.

Estas consideraciones sugieren la siguiente generalización que incluye a rectas y planos como casos especiales. Un subconjunto $\mathcal{S}$ de un espacio vectorial $\mathcal{U}$ se llama un **subespacio lineal** (o, a menudo, simplemente **subespacio**) si toda combinación lineal de elementos de $\mathcal{S}$ está en $\mathcal{S}$.

Ejemplo 2. Damos algunos ejemplos de subespacios.

(a) El conjunto de todos los vectores en $\mathcal{R}^n$ cuya primera coordenada es 0 es un subespacio, ya que toda combinación lineal de tales vectores tendrá primera coordenada 0.

(b) Si $\mathcal{U}$ es un espacio vectorial, $\mathcal{U}$ es un subespacio. A veces se usa el término subespacio **propio** para indicar un subespacio que no es todo el espacio. En todo espacio vectorial, el vector cero forma un subespacio. Este subespacio se llama **trivial** (porque lo es).

(c) Si $\mathcal{U} \xrightarrow{f} \mathcal{W}$ es una función lineal, el conjunto $\mathfrak{N}$ de todos los vectores $\mathbf{x}$ de $\mathcal{U}$ tales que $f(\mathbf{x}) = 0$ es un subespacio de $\mathcal{U}$ llamado **núcleo** de f. Si $\mathbf{x}_1, \ldots, \mathbf{x}_k$ están en $\mathfrak{N}$ y $\mathbf{x} = \sum_{i=1}^{k} r_i\mathbf{x}_i$, entonces $f(\mathbf{x}) = \sum_{i=1}^{k} r_i f(\mathbf{x}_i) = 0$, ya que f es lineal y cada $f(\mathbf{x}_i)$ es 0. Por tanto, $\mathbf{x}$ también está en $\mathfrak{N}$ y $\mathfrak{N}$ es un subespacio de $\mathcal{U}$. En particular, el conjunto de los vectores (x, y, z) de $\mathcal{R}^3$ tales que

$$x + 2y + 3z = 0,$$

o bien

$$(1 \quad 2 \quad 3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

es un subespacio de $\mathcal{R}^3$, ya que es el núcleo de la función lineal de $\mathcal{R}^3$ en $\mathcal{R}^2$ definida por la matriz mencionada.

(d) La imagen de una función lineal f definida en un espacio, es un subespacio del espacio imagen de f. La razón para esto es que si $\mathbf{y}_1, \ldots, \mathbf{y}_k$ están en la imagen de f, esto significa que existen vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ en el dominio de f tales que $f(\mathbf{x}_i) = \mathbf{y}_i$, para $i = 1, \ldots, k$. Pero, entonces, por la linealidad de f, una combinación lineal de los $\mathbf{y}_i$ tiene la forma

$$\begin{aligned} r_1\mathbf{y}_1 + \ldots + r_k\mathbf{y}_k &= r_1 f(\mathbf{x}_1) + \ldots + r_k f(\mathbf{x}_k) \\ &= f(r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k). \end{aligned}$$

Como el dominio de f es un espacio vectorial, $r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k$ es un vector en este espacio y, por tanto, $r_1\mathbf{y}_1 + \ldots + r_k\mathbf{y}_k$ está en la imagen de f. En particular, el conjunto de todos los vectores en $\mathcal{R}^3$ de la forma

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

es un subespacio porque es el rango de la función lineal definida por la matriz 3-por-2, indicada anteriormente.

Definimos el **subespacio generado** por un conjunto S de vectores como el conjunto de todas las combinaciones lineales de elementos de S. Es fácil comprobar que éste es un subespacio. En efecto, sea $\mathbf{x} = \sum_{i=1}^{k} r_i\mathbf{x}_i$ una combinación lineal de vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ en el subespacio generado por S. Entonces $\mathbf{x}_i = \sum_{j=1}^{n} s_{ij}\mathbf{u}_j$ para ciertos vectores $\mathbf{u}_j$ en S. Entonces, $\mathbf{x} = \sum_{i=1}^{k} r_i\left(\sum_{j=1}^{n} s_{ij}\mathbf{u}_j\right) = \sum_{j=1}^{n} t_j\mathbf{u}_j$, donde $t_j = \sum_{i=1}^{k} r_i s_{ij}$. Esto muestra que $\mathbf{x}$ está en el subespacio generado por S.

Recordemos la definición general de un espacio vectorial (dada en la Sec. 1) como un conjunto $\mathcal{U}$ con operaciones de suma y multiplicación numérica tal que para elementos $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ de $\mathcal{U}$ y números reales r, s cualesquiera,

1. $r\mathbf{x} + s\mathbf{x} = (r + s)\mathbf{x}$.
2. $r\mathbf{x} + r\mathbf{y} = r(\mathbf{x} + \mathbf{y})$.
3. $r(s\mathbf{x}) = (rs)\mathbf{x}$.
4. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
5. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
6. Existe un elemento 0 en $\mathcal{U}$ tal que $\mathbf{x} + 0 = \mathbf{x}$.
7. Para todo $\mathbf{x}$ en $\mathcal{U}$, $\mathbf{x} + (-1)\mathbf{x} = 0$.

Ahora, si $\mathcal{S}$ es un subespacio de un espacio vectorial $\mathcal{U}$, entonces las operaciones de suma y multiplicación numérica dada para $\mathcal{U}$, dan siempre un resultado en $\mathcal{S}$ cuando se aplican a elementos de $\mathcal{S}$, puesto que $\mathbf{x} + \mathbf{y} = 1\mathbf{x} + 1\mathbf{y}$ y $r\mathbf{x} = r\mathbf{x} + 0\mathbf{y}$ son combinaciones lineales de $\mathbf{x}$ e $\mathbf{y}$. Las leyes 1 a 5 son ciertamente verdaderas en $\mathcal{S}$, ya que son válidas para todos los elementos de $\mathcal{U}$. El vector cero pertenece a $\mathcal{S}$, ya que $0 = 0\mathbf{x}$ para cualquier $\mathbf{x}$ en $\mathcal{S}$ también, si $\mathbf{x}$ está en $\mathcal{S}$ entonces $(-1)\mathbf{x} = -\mathbf{x}$ también está en $\mathcal{S}$. Por tanto, 6 y 7 también son válidas. En otras palabras, hemos demostrado

4.1 Teorema

Todo subespacio de un espacio vectorial es un espacio vectorial, con las operaciones heredadas del espacio original.

A pesar de que los subespacios de $\mathcal{R}^n$ son ejemplos importantes de espacios vectoriales, hay muchos otros. A continuación daremos algunos.

Ejemplo 3. (a) Sea $\mathcal{U}$ el conjunto de todas las funciones continuas de una variable real y con valores reales. Definimos $f + g$ y rf de la manera obvia como las funciones cuyos valores para un número x cualquiera son $f(x) + g(x)$ y $rf(x)$, respectivamente. (Por supuesto, estamos usando los teoremas que dicen que $f + g$ y rf son continuas si f y g lo son.) Es fácil ver que las leyes de los espacios vectoriales se satisfacen.

(b) Sea P el subespacio de $\mathcal{U}$ formado por todos los polinomios, es decir, de todas las funciones f que pueden ser expresadas por una fórmula $f(x) = a_0 + a_1x + \ldots + a_kx^k$ para algunas constantes $a_0, \ldots, a_k$. (¿Qué hay que comprobar para verificar que éste es un subespacio?)

(c) Sea P_m el subespacio de los polinomios de grado menor o igual a n, es decir, aquellos que no requieren potencias de x mayores que n para su expresión. Para $k \leq n$, P_k es un subespacio de P_n y todos los P_n son subespacios de P.

(d) Sea $C^{(k)}$ el espacio vectorial de las funciones reales $f(x)$ cuyas n primeras derivadas son continuas para todo x real. Entonces $C^{(k+1)}$ es un subespacio de $C^{(k)}$. Si denotamos por $C^{(\infty)}$ el espacio vectorial de funciones que tienen derivadas de *todo* orden, entonces $C^{(\infty)}$ es un subespacio de $C^{(k)}$ para todo k.

Es habitual decir que las rectas, los planos y el espacio ordinario tienen dimensiones 1, 2 y 3, respectivamente. Es posible definir la dimensión de cualquier espacio vectorial. Los ejemplos de rectas y planos sugieren que el subespacio generado por k vectores debe tener dimensión k. Esto no es totalmente adecuado, ya que, por ejemplo, el subespacio generado por dos vectores colineales es una recta y no un plano. Para hacer esto de una manera adecuada, necesitamos la noción de independencia lineal, introducida en el Cap. 1. Recordamos la siguiente definición.

Un conjunto $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ de vectores es **(linealmente) independiente** si el único conjunto $r_1, \ldots, r_k$ de números tales que $r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = 0$ es el conjunto $r_1 = r_2 = \ldots = r_k = 0$. Un conjunto de vectores es **(linealmente) dependiente** si no es independiente. Para que $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ sea dependiente deben existir números $r_1, \ldots, r_k$, *no* todos cero, tales que $r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = 0$.

Ejemplo 4. (a) Los cuatro vectores $\mathbf{x}_1 = (2, 0, 0)$, $\mathbf{x}_2 = (0, -2, 0)$, $\mathbf{x}_3 = (0, 0, 3)$, $\mathbf{x}_4 = (2, -2, 3)$ son linealmente dependientes, ya que $\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4 = 0$. El conjunto formado por los tres vectores $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ es independiente, ya que $r\mathbf{x}_1 + s\mathbf{x}_2 + t\mathbf{x}_3 = 0$ si, y solo si $r = s = t = 0$.

(b) El conjunto de dos vectores $\mathbf{x}$ e $\mathbf{y}$ es independiente solo si ninguno

es múltiplo numérico del otro. Por ejemplo, si $\mathbf{y} = 3\mathbf{x}$, entonces $3\mathbf{x} - \mathbf{y} = 0$ y los vectores son dependientes.

Un conjunto de vectores linealmente independiente que genere un espacio $\mathcal{U}$ se llama una **base** de $\mathcal{U}$.

Ejemplo 5. (a) La base natural de vectores $\mathbf{e}_1, \ldots, \mathbf{e}_n$, donde $\mathbf{e}_i$ tiene coordenada i igual a 1 y todas las otras 0, forman una base de $\mathcal{R}^n$. La confirmación de que las $\mathbf{e}_i$ son linealmente independientes y que generan a $\mathcal{R}^n$, es dejado al lector.

(b) En el espacio P_n del Ejemplo 3(c), los polinomios $\mathbf{x}_0 = 1$, $\mathbf{x}_1 = 1, \ldots, \mathbf{x}_n = x^n$ forman una base. Es obvio que si $f(x) = a_0 + a_1x + \ldots + a_nx^n$ es un polinomio de grado menor o igual que n, es una combinación lineal $a_0\mathbf{x}_0 + \ldots + a_n\mathbf{x}_n$ de los $\mathbf{x}$. Si $a_0\mathbf{x}_0 + \ldots + a_n\mathbf{x}_n$ es la función cero, entonces (ya que un polinomio de grado menor o igual que n no puede tener más de n raíces, a no ser que todos los coeficientes sean cero) $a_0 = a_1 = \ldots = a_n = 0$.

A pesar de que es verdad que todo espacio vectorial tiene una base, generalmente, consideraremos solo aquellos que tienen una base con un número finito de elementos. El teorema siguiente implica que si un espacio es generado por un conjunto finito (posiblemente linealmente dependiente), entonces tiene una base finita.

4.2 Teorema

Sea $\mathcal{U}$ el espacio generado por $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Entonces, $\mathcal{U}$ contiene solo al vector cero o un subconjunto de los $\mathbf{x}$ es una base para $\mathcal{U}$.

Demostración. Si el conjunto $\mathbf{x}_1, \ldots, \mathbf{x}_n$ es linealmente independiente, él es una base. En caso contrario, existe una relación $r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = 0$, con algún r, digamos r_k, distinto de 0. Entonces, podemos dividir por r_k y obtener $\mathbf{x}_k = -(r_1/r_k)\mathbf{x}_1 - \ldots - (r_{n-1}/r_k)\mathbf{x}_n$, donde $\mathbf{x}_k$ no aparece en el término de la derecha. Si sustituimos esta expresión para $\mathbf{x}_k$ en cualquier combinación lineal de todos los $\mathbf{x}$, se obtiene una combinación lineal que no involucra a $\mathbf{x}_k$. En otras palabras, si se quita a $\mathbf{x}_k$, el espacio generado por los $\mathbf{x}$ restantes sigue siendo $\mathcal{U}$. Si el conjunto resultante no es independiente, se puede repetir este proceso. Esto debe terminar en un número finito de pasos ya sea porque se ha obtenido una base o porque se han eliminado todos los vectores. Esta última posibilidad no puede ocurrir si el espacio tiene algún vector no cero.

El teorema siguiente es el paso más importante en el desarrollo de la teoría de la dimensión.

4.3 Teorema

Si un espacio vectorial es generado por n vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ y $\mathbf{y}_1, \ldots, \mathbf{y}_k$ es un conjunto de k vectores linealmente independientes en el espacio, entonces $k \leq n$.

Demostración. Es fácil demostrar la afirmación si $n = 1$. En este caso, el espacio generado por $\mathbf{x}_1$ consiste en todos los múltiplos numéricos de $\mathbf{x}_1$. Entonces, si $\mathbf{y}_1$ e $\mathbf{y}_2$ son dos vectores en el espacio, uno de ellos debe ser un múltiplo del otro y $\mathbf{y}_1$ e $\mathbf{y}_2$ no pueden ser linealmente independientes. Procedemos por inducción y suponemos que en cualquier espacio generado por $n - 1$ vectores, no puede haber un conjunto linealmente independiente con más de $n - 1$ vectores. Ahora, supongamos dado un espacio generado por n vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ y sean $\mathbf{y}_1, \ldots, \mathbf{y}_k$ k vectores, donde $k > n$. Si probamos que los $\mathbf{y}$ son linealmente dependientes, habremos demostrado que la afirmación del teorema es válida para un conjunto generador de n elementos y la demostración por inducción estará terminada. Cada uno de los vectores $\mathbf{y}_1, \ldots, \mathbf{y}_{n+1}$ puede ser expresado como una combinación lineal de $\mathbf{x}_1, \ldots, \mathbf{x}_n$, lo que significa que existen números a_{ij} tales que $\mathbf{y}_i = \sum_{j=1}^{n} a_{ij}\mathbf{x}_j$, para $i = 1, \ldots, n + 1$. Si los $n + 1$ números a_{i1}, son todos cero, entonces todos los $\mathbf{y}$ están en el espacio generado por los $n - 1$ vectores $\mathbf{x}_2, \ldots, \mathbf{x}_n$ y, por la hipótesis de inducción, los $\mathbf{y}$ deben ser dependientes, como queríamos demostrar. En caso contrario (renumerando si es preciso), podemos suponer que a_{11} es distinto de cero. Entonces definimos n vectores $\mathbf{z}_2, \ldots, \mathbf{z}_{n+1}$ por $\mathbf{z}_i = \mathbf{y}_i - a_{11}^{-1}a_{i1}\mathbf{y}_1$. Usando las ecuaciones que expresan $\mathbf{y}_i$ en términos de los $\mathbf{x}$, es fácil ver que

$$\mathbf{z}_i = \sum_{j=2}^{n}(a_{ij} - a_{11}^{-1}a_{i1}a_{1j})\mathbf{x}_j,$$

de modo que los $\mathbf{z}$ son combinaciones lineales de los $n - 1$ vectores $\mathbf{x}_2, \ldots, \mathbf{x}_n$. Por la hipótesis de inducción, existen números $r_2, \ldots, r_{n+1}$, no todos ceros, tales que $r_2\mathbf{z}_2 + \ldots + r_{n+1}\mathbf{z}_{n+1} = 0$. Usando la definición de los $\mathbf{z}$, esta última relación se transforma en

$$-a_{11}^{-1}(r_2a_{21} + \ldots + r_{n+1}a_{n+11})\mathbf{y}_1 + r_2\mathbf{y}_2 + \ldots + r_{n+1}\mathbf{y}_{n+1} = 0.$$

Pero como no todos los r son cero, tenemos que $\mathbf{y}_1, \ldots, \mathbf{y}_{n+1}$ son dependientes, como queríamos demostrar.

4.4 Teorema

Sea $\mathcal{U}$ un espacio vectorial con una base con n elementos. Entonces, toda otra base de $\mathcal{U}$ tiene n elementos.

Demostración. Sean $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ dos bases de $\mathcal{U}$. Como ambos conjuntos son independientes y ambos son conjuntos generadores, el Teorema 4.3 dice que $k \leq n$ y $n \leq k$.

La **dimensión** de un espacio vectorial que tiene un conjunto finito de generadores es el número de elementos en cualquier base del espacio. (La dimensión del espacio vectorial que solo contiene al vector cero es, por definición, 0.) Denotamos por dim ($\mathcal{U}$) la dimensión de $\mathcal{U}$. Nótese que el Teorema 4.2 garantiza la existencia de una base y el Teorema 4.4 garantiza que la dimensión no depende de la base elegida.

Ejemplo 6. (a) Por el Ejemplo 5(a), dim $(\mathcal{R}^n) = n$.

(b) Por el Ejemplo 5(b), dim $(\mathcal{R}_n) = n + 1$.

(c) El espacio P de todos los polinomios (Ejemplo 4(b)) no tiene un conjunto finito de generadores. Si tuviera uno con k elementos, entonces, como $1, x, x^2, \ldots, x^k$ son $k + 1$ elementos linealmente independientes de P, tendríamos una contradicción al Teorema 4.3.

Un espacio vectorial con una base finita se llama de **dimensión finita.** Como lo indica el Ejemplo 6(c), existen espacios que no son de dimensión finita. El Teorema 4.2 afirma que podemos obtener una base a partir de un conjunto finito de generadores simplemente eliminando algunos de ellos. El teorema siguiente dice que en un espacio de dimensión finita se puede obtener una base a partir de un conjunto linealmente independiente agregando más vectores.

4.5 Teorema

Sea $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ un conjunto linealmente independiente en un espacio $\mathcal{U}$. Si S no es una base, entonces S puede ser extendido a una base finita de $\mathcal{U}$ o bien puede ser extendido a una sucesión infinita de vectores independientes, en cuyo caso, $\mathcal{U}$ no es de dimensión finita.

Demostración. Supongamos que $\mathbf{x}_1, \ldots, \mathbf{x}_k$ son linealmente independientes y que no generan a $\mathcal{U}$. Entonces existe un vector $\mathbf{y}$ que no es una combinación lineal de $\mathbf{x}_1, \ldots, \mathbf{x}_k$. Sea $\mathbf{x}_{k+1} = \mathbf{y}$. Afirmamos que el conjunto $\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1}$ es linealmente independiente. Supongamos $r_1\mathbf{x}_1 + \ldots + r_{k+1}\mathbf{x}_{k+1} = 0$. Debemos demostrar que todos los r son 0. Si r_{k+1} fuera distinto de cero, podríamos escribir $\mathbf{x}_{k+1} = -(r_1/r_{k+1})\mathbf{x}_1 - \ldots - (r_k/r_{k+1})\mathbf{x}_k$, lo cual es imposible, ya que $\mathbf{x}_{k+1}$ no es combinación lineal de los otros $\mathbf{x}$. Por tanto, $r_{k+1} = 0$ y $r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = 0$. Como $\mathbf{x}_1, \ldots, \mathbf{x}_k$ son linealmente independientes, esto implica que $r_1 = \ldots = r_k = 0$. En otras palabras, si un conjunto linealmente independiente no genera todo el espacio, entonces se le puede agregar otro vector tal que el conjunto resultante también es independiente.

Este proceso de agregar vectores puede ser continuado, a no ser que se llegue a un conjunto de generadores. Si se llega a un conjunto de generadores, entonces tenemos una base para $\mathcal{U}$ y éste es de dimensión finita. En caso contrario, se puede encontrar un conjunto independiente arbitrariamente grande y $\mathcal{U}$ no puede ser de dimensión finita. Esto completa la demostración.

4.6 Teorema

Si $\mathcal{S}$ es un subespacio del espacio vectorial de dimensión finita $\mathcal{U}$ tal que dim $(\mathcal{U}) = n$, entonces $\mathcal{S}$ es de dimensión finita y dim $(\mathcal{S}) \leq n$. Toda base de $\mathcal{S}$ puede ser extendida a una base de $\mathcal{U}$ y si dim $(\mathcal{S}) = n$, entonces $\mathcal{S} = \mathcal{U}$.

Demostración. Si $\mathcal{S}$ solo contiene a 0, entonces tiene dimensión 0. En caso contrario, aplicamos el Teorema 4.5 a un vector en $\mathcal{S}$ distinto de cero. Como ningún subconjunto independiente de $\mathcal{U}$ puede tener más de n elementos, esta

construcción nos proporciona una base finita para $\mathcal{S}$. Esta base es un conjunto linealmente independiente de $\mathcal{U}$ que, si es necesario, puede ser extendida a una base de $\mathcal{U}$ (de nuevo por el Teorema 4.5). Tenemos así una base para $\mathcal{S}$ que es un subconjunto de una base para $\mathcal{U}$. Si dim $(\mathcal{S}) = n$, entonces este subconjunto debe ser igual al conjunto completo, y entonces es una base de $\mathcal{S}$ y de $\mathcal{U}$ y $\mathcal{S} = \mathcal{U}$.

El siguiente teorema expresa una propiedad importante de las funciones lineales:

4.7 Teorema

Sea f una función lineal definida en un espacio vectorial de dimensión finita. Entonces

$$\dim (\text{núcleo de } f) + \dim (\text{imagen de } f) = \dim (\text{dominio de } f).$$

Demostración. Como en todo teorema acerca de dimensiones, para demostrarlo, debemos encontrar bases adecuadas para los espacios involucrados. Denotemos por $\mathcal{U}$ el dominio de f, por $\mathfrak{N}$ a su núcleo y por $\mathcal{W}$ a su imagen. Sea $\mathbf{v}_1, \ldots, \mathbf{v}_k$ una base de $\mathfrak{N}$ que extendemos a una base $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_k$ de $\mathcal{U}$. (El Teorema 4.6 garantiza esta construcción.) Entonces dim $(\mathfrak{N}) = k$, y dim $(\mathcal{U}) = k + r$. Sean $\mathbf{w}_1 = f(\mathbf{u}_1), \ldots, \mathbf{w}_2 = f(\mathbf{u}_r)$. Afirmamos que $\mathbf{w}_1, \ldots, \mathbf{w}_r$ es una base para $\mathcal{W}$, lo cual implica que dim $(\mathcal{W}) = r_1$, lo que demuestra el teorema. Es obvio que los vectores $\mathbf{w}_1, \ldots, \mathbf{w}_r$ generan $\mathcal{W}$, ya que si $\mathbf{y} = f(\mathbf{x})$ es un vector cualquiera en la imagen de f, podemos escribir

$$\mathbf{x} = \sum_{i=1}^{k} a_i\mathbf{v}_i + \sum_{i=1}^{r} b_i\mathbf{u}_i,$$

y entonces

$$\mathbf{y} = \sum_{i=1}^{k} a_i f(\mathbf{v}_i) + \sum_{i=1}^{r} b_i f(\mathbf{u}_i) = 0 + \sum_{i=1}^{r} b_i\mathbf{w}_i,$$

lo cual muestra que $\mathbf{y}$ es una combinación lineal de los $\mathbf{w}$. Falta demostrar que $\mathbf{w}_1, \ldots, \mathbf{w}_r$ son linealmente independientes. Supongamos que $\sum_{i=1}^{r} b_i\mathbf{w}_i = 0$. Esto significa que el vector $\sum_{i=1}^{r} b_i\mathbf{u}_i$ está en el núcleo de f y es, por tanto, una combinación lineal $\sum_{i=1}^{r} a_i\mathbf{v}_i$, de modo que $a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k - b_1\mathbf{u}_1 - \ldots - b_r\mathbf{u}_r = 0$. Como $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_r$ son linealmente independientes, esto es posible solo si todos los b (y todos los a) son cero, lo cual demuestra que los $\mathbf{w}$ son linealmente independientes.

Ejemplo 7. Supongamos que $m \leq n$ y definimos $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$, de modo que $f(\mathbf{x})$ sea el vector columna de dimensión m formado por las primeras m coordenadas de $\mathbf{x}$. La imagen de f es todo $\mathcal{R}^m$ y su núcleo, de dimensión $n - m$, consiste en todos los vectores de $\mathcal{R}^n$ cuyas primeras m componentes son todas cero. Las definiciones de recta y plano pueden ser unificadas en términos de subespacios de $\mathcal{R}^n$. Como un subespacio siempre contiene al vector cero, de-

finimos una **recta por el origen** como un subespacio de dimensión 1 y un **plano por el origen** como un subespacio de dimensión 2. La Fig. 5(a) muestra ejemplos en $\mathcal{R}^3$. Una recta o un plano que no pasa por el origen se puede obtener trasladando uno que sí pase por el origen, es decir, sumando un vector fijo a todo vector en el subespacio. Decimos que un subconjunto $\mathcal{S}$ de un espacio vectorial es un **subespacio afín** si él puede ser expresado en la forma

$$\mathcal{U} + \mathbf{b},$$

donde $\mathcal{U}$ es un subespacio lineal y $\mathbf{b}$ un vector. La dimensión del espacio afín es por definición, la dimensión de $\mathcal{U}$. Un subespacio afín de dimensión 1 se llama una **recta** y uno de dimensión 2, un **plano**. La Fig. 5(b) muestra ejemplos en $\mathcal{R}^3$. Dos subespacios afines son **paralelos** si ellos son trasladados uno del otro.

Dos de las maneras más importantes de describir subespacios, a saber, la imagen y el núcleo de una función lineal, dan maneras estándar para describir rectas y planos. La representación paramétrica de planos, discutida en la Sección 2 del Cap. 1, se obtiene partiendo de una función lineal $\mathcal{R}^2 \xrightarrow{L} \mathcal{R}^3$ con imagen de dimensión 2 con base $\mathbf{x}_1, \mathbf{x}_2$. Escribimos

$$L(u, v) = u\mathbf{x}_1 + v\mathbf{x}_2$$

y entonces formamos la función afín

$$A(u, v) = u\mathbf{x}_1 + v\mathbf{x}_2 + \mathbf{x}_0$$

sumando un vector fijo $\mathbf{x}_0$. La representación de una recta tiene la forma similar

$$A(t) = t\mathbf{x} + \mathbf{x}_0,$$

en la cual los vectores $t\mathbf{x}_1$ forman un subespacio de dimensión 1 que es la imagen de la función lineal $\mathcal{R} \xrightarrow{L} \mathcal{R}^3$.

Para representar una recta o plano por el origen como el núcleo de una función lineal, consideremos todos los vectores $\mathbf{x}$ solución de una ecuación *homogénea*

$$L(\mathbf{x}) = 0,$$

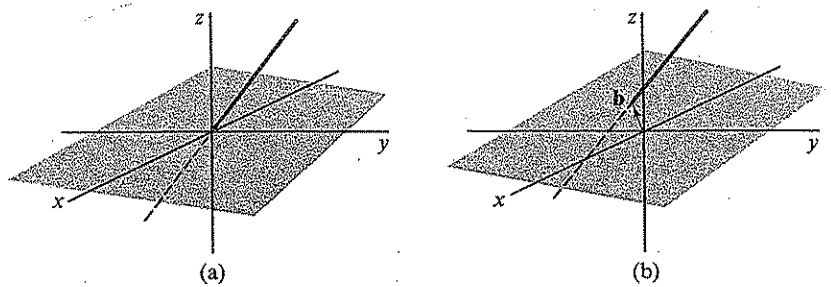


Figura 5

donde L es lineal. Un subespacio afín se obtiene sumando un vector fijo $\mathbf{x}_0$ al resultado 0, lo que es lo mismo, resolviendo la ecuación

$$L(\mathbf{x} - \mathbf{x}_0) = 0.$$

Como L es lineal, esta última ecuación puede ser escrita como una ecuación *no homogénea*

$$L(\mathbf{x}) = L(\mathbf{x}_0).$$

En secciones anteriores de este capítulo se han discutido métodos para resolver este tipo de ecuaciones. Si $\mathbf{x} = (x, y)$ es de dimensión 2, entonces la ecuación (1) toma la forma

$$ax + by = c,$$

la cual es la forma habitual de representar una recta. Si $\mathbf{x} = (x, y, z)$ es de dimensión 3, la ecuación

$$a_1x + b_1y + c_1z = d_1$$

tiene como solución un plano en $\mathcal{R}^3$ mientras que un sistema de dos ecuaciones

$$a_1x + b_1y + c_1z = d_1$$
$$a_2x + b_2y + c_2z = d_2$$

tiene como solución la intersección de dos planos, es decir, una recta o un plano.

Ejemplo 8. Si tratamos de resolver las ecuaciones

$$x - y + 2z = 1$$
$$x + y + 3z = 0$$

mediante operaciones en sus filas, sumamos la segunda ecuación a la primera, obteniendo

$$2x \qquad + 5z = 1$$
$$x + y + 3z = 0.$$

Multiplicando la primera ecuación por $\frac{1}{2}$ y restando a la segunda, se obtiene

$$x \qquad + \tfrac{5}{2}z = \tfrac{1}{2}$$
$$y + \tfrac{1}{2}z = -\tfrac{1}{2}.$$

Podemos representar paramétricamente al conjunto de todas las soluciones de este par de ecuaciones como un subespacio afín haciendo $z = t$. Obtenemos

$$x = -\tfrac{5}{2}t + \tfrac{1}{2}$$
$$y = -\tfrac{1}{2}t - \tfrac{1}{2}$$
$$z = t$$

o bien

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} -\frac{5}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix}.$$

Tenemos, entonces, una recta de soluciones que se puede describir partiendo de una recta por el origen con la dimensión $\mathbf{x}_1 = (-\frac{5}{2}, -\frac{1}{2}, 1)$ y después trasladándola sumando el vector $\mathbf{x}_0 = (\frac{1}{2}, -\frac{1}{2}, 0)$. La Fig. 6 muestra el resultado.

Si tenemos un sistema de tres ecuaciones, su conjunto de soluciones es la intersección de los conjuntos de soluciones de cada una de las ecuaciones. La situación habitual es que los tres planos se corten en un punto. Si cambiamos los términos de la derecha de las ecuaciones, traslada paralelamente los planos, y ellos seguirán cortándose en un punto. Otra posibilidad es que los planos sean paralelos. El sistema homogéneo asociado tendrá un conjunto de soluciones de dimensión 2, si los tres planos coinciden, el sistema no homogéneo también tendrá un conjunto de dimensión 2. En caso contrario, no hay solución. Otra posibilidad (que no tiene una analogía en dos dimensiones) es que no hay dos planos paralelos y los planos que representan las soluciones de las ecuaciones *homogéneas* se cortan en una recta común. Entonces (a no ser que todos pasen por una recta), cada par de planos formados por las soluciones de las ecuaciones *no* homogéneas se cortarán en una recta paralela al tercer plano y no habrá solución común para las tres ecuaciones.

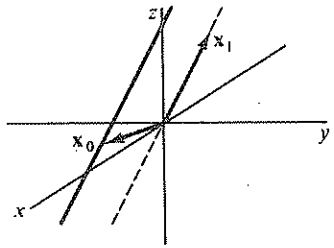


Figura 6

EJERCICIOS

1. ¿Cuáles de los siguientes subconjuntos de $\mathcal{R}^3$ son subespacios? En cada caso demuestre que el subconjunto es un subespacio o encuentre una combinación lineal de elementos del subconjunto que no se encuentre en él.

   (a) Todos los vectores (x_1, x_2, x_3) con $x_1 + x_2 = 0$.
   (b) Todos los vectores con $x_3 = 0$.
   (c) Todos los vectores que satisfacen (a) y (b).
   (d) Todos los vectores que satisfacen (a) o (b).
   (e) Todos los vectores con $x_1 = (x_2)^3$.
   (f) Todos los vectores que satisfacen (a) y (e).

2. Sean $\mathbf{x}_1 = (1, 2, 3)$, $\mathbf{x}_2 = (-1, 2, 1)$, $\mathbf{x}_3 = (1, 1)$, $\mathbf{x}_4 = (1, 1, 0)$.
   (a) Demuestre que $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ es un conjunto linealmente dependiente. (Resuelva un sistema apropiado de ecuaciones.)

   (b) Exprese $\mathbf{x}_1$ como combinación lineal de $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ por un método similar al usado en la parte (a).

   [*Resp.* $\mathbf{x}_1 = \frac{1}{3}\mathbf{x}_2 + \frac{8}{3}\mathbf{x}_3 - \frac{4}{3}\mathbf{x}_4$.]

3. Sea $C[a, b]$ el espacio vectorial de las funciones reales continuas definidas en el intervalo $[a, b]$. Sea $C_0[a, b]$ el subconjunto de las funciones en $C[a, b]$ tales que $f(a) = f(b) = 0$.

   (a) Demuestre que $C_0[a, b]$ es un subespacio propio de $C[a, b]$.
   (b) ¿Qué sucede si la condición $f(a) = f(b) = 0$ se remplaza por $f(a) = 1$, $f(b) = 0$?

4. Demuestre que la intersección de dos subespacios es un subespacio.

5. El Ejercicio 1(d) demuestra que la unión de dos subespacios no es necesariamente un subespacio. Demuestre que la unión de dos subespacios es un subespacio si, y solo si, uno de ellos está contenido en el otro.

6. Demuestre que la imagen de una función lineal puede ser un subespacio propio del espacio imagen.

7. Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ la función definida por la multiplicación por la matriz m-por-n, A. Demuestre que su imagen está generada por las columnas de A (consideradas como elementos de $\mathcal{R}^m$).

8. Dados dos subconjuntos $\mathcal{A}$ y $\mathcal{B}$ de un espacio vectorial, sea $\mathcal{A} + \mathcal{B}$ el conjunto de todos los vectores expresables como una suma $\mathbf{a} + \mathbf{b}$ con $\mathbf{a}$ en $\mathcal{A}$ y $\mathbf{b}$ en $\mathcal{B}$. Demuestre que si $\mathcal{A}$ y $\mathcal{B}$ son subespacios entonces $\mathcal{A} + \mathcal{B}$ también lo es.

9. Demuestre que si 0 es el único elemento en la intersección de dos subespacios $\mathcal{S}$ y $\mathcal{T}$ entonces

$$\dim (\mathcal{S} + \mathcal{T}) = \dim (\mathcal{S}) + \dim (\mathcal{T})$$

   [*Indicación*: Demuestre que la unión de una base para $\mathcal{S}$ y una base para $\mathcal{T}$ es una base para $\mathcal{S} + \mathcal{T}$].

10. Sean $\mathcal{S}$ y $\mathcal{T}$ subespacios cualesquiera. Demuestre que

$$\dim (\mathcal{S} + \mathcal{T}) = \dim (\mathcal{S}) + \dim (\mathcal{T}) - \dim (\mathcal{S} \cap \mathcal{T})$$

    [*Indicación*: Parta de una base para $\mathcal{S} \cap \mathcal{T}$ y extiéndala a una base para $\mathcal{S}$ y a una para $\mathcal{T}$ (Teorema 4.6)].

11. Sea $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}^2$ definida por la matriz

$$\begin{pmatrix} 1 & -3 & 2 \\ -2 & 6 & -4 \end{pmatrix}.$$

    Encuentre una base para el núcleo de f y una para la imagen de f. Verifique que se cumple el Teorema 4.7.

12. Describa el conjunto de soluciones de las ecuaciones siguientes como subespacios

afines, es decir, especifique una base del subespacio vectorial y las coordenadas de un vector con el cual hay que transladarlo para obtener el subespacio afín.

(a) $x + y = 1$
(b) $2x + y = 1$
$2x - 3y + z = 2.$
(c) $x + 2y + 3z = 10$
$4x + 5y + 6z = 11$
$7x + 8y + 9z = 12.$ [*Resp*: $t(1, -2, 1) + (0, -9, \frac{28}{3})$.]
(d) $x - y + z = 2.$

13. Para cada uno de los siguientes conjuntos de tres ecuaciones con tres incógnitas, determine cuál de las posibilidades geométricas consideradas al final de la sección es válida.

(a) $x + 2y - 3z = 2$
$3x + 6y - 9z = 0$
$-2x - 4y + 6z = 3.$
(b) $x + y - z = 1$
$-x + y + z = 3$
$x - y + z = 5.$
(c) $x - y + z = 0$
$-x + 2y - 3z = 2$
$2x - y = 6.$

5. Funciones lineales

En el Cap. 1 vimos que las matrices obedecen algunas de las reglas de la suma y multiplicación de números. Es más, hemos visto que, dada una función lineal f de $\mathcal{R}^n$ en $\mathcal{R}^m$, existe una matriz m-por-n. A tal que $f(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ en $\mathcal{R}^n$. Usando estos hechos, podríamos demostrar que las funciones lineales de $\mathcal{R}^n$ en $\mathcal{R}^m$ obedecen leyes algebraicas del mismo modo que las matrices. Sin embargo, sucede que estas reglas se aplican a una clase más grande de funciones y no solo a las que son representadas por matrices. Por esta razón, demostraremos estas reglas en su forma general, para después aplicarlas en la Sec. 6 a un análisis sistemático de algunas ecuaciones diferenciales lineales. Empezamos por describir las operaciones de suma, multiplicación escalar y composición de funciones. Sean f y g dos funciones con el mismo dominio y el mismo espacio vectorial como espacio imagen. Entonces la función $f + g$ es la **suma** de f y g y está definida por

$$(f+g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

para todo $\mathbf{x}$ en el dominio de f y de g. En forma análoga, si r es un número, entonces rf es el **múltiplo numérico** de f por r y está definido por

$$rf(\mathbf{x}) = r(f(\mathbf{x})).$$

En el Cap. 1 definimos la composición de funciones, pero repetiremos la definición. Exigimos ahora que la imagen de f esté contenida en el dominio de g. Entonces $g \circ f$ es la **composición** de f y g y está definida por

$$g \circ f(\mathbf{x}) = g(f(\mathbf{x})).$$

Un ejercicio fácil es demostrar que las sumas, múltiplos numéricos y composición de funciones lineales son lineales.

Ejemplo 1. Supongamos que $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^2$ y $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^2$ están dadas por

$$f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+y\\x-y\end{pmatrix} = \begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

y

$$g\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}2x+y\\x+3y\end{pmatrix} = \begin{pmatrix}2 & 1\\1 & 3\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Entonces

$$(f+g)\begin{pmatrix}x\\y\end{pmatrix} = f\begin{pmatrix}x\\y\end{pmatrix} + g\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3x+2y\\2x+2y\end{pmatrix} = \begin{pmatrix}3 & 2\\2 & 2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

También, $3f$ está definida por

$$3f\begin{pmatrix}x\\y\end{pmatrix} = 3\left(f\begin{pmatrix}x\\y\end{pmatrix}\right) = \begin{pmatrix}3x+3y\\3x-3y\end{pmatrix} = \begin{pmatrix}3 & 3\\3 & -1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Por último, de acuerdo con el Teorema 4.4 del Cap. 1, $g \circ f$ está dada por la multiplicación matricial, de modo que

$$g \circ f\begin{pmatrix}x\\y\end{pmatrix} = g\left(f\begin{pmatrix}x\\y\end{pmatrix}\right) = \begin{pmatrix}2 & 1\\1 & 3\end{pmatrix}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3 & 1\\4 & -2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Nótese que en el ejemplo anterior la matriz de $f + g$ es igual a la matriz de f sumada a la matriz de g y que la matriz de $3f$ es 3 veces la matriz de f. Además, en el Teorema 4.5 del Cap. 1, demostramos que la matriz de la com-

puesta $g \circ f$ es el producto de las matrices de g y f en el mismo orden. El resultado general es el siguiente.

5.1 Teorema

Sean f y g funciones lineales de $\mathfrak{R}^n$ en $\mathfrak{R}^m$ con matrices A y B, respectivamente. Entonces

(i) La matriz de $f + g$ es $A + B$.
(ii) Para todo número r, la matriz de rf es rA.
(iii) Si $g \circ f$ está definida, su matriz es BA.

Demostración. Para demostrar (i), escribimos:

$$(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$
$$= A\mathbf{x} + B\mathbf{x} = (A + B)\mathbf{x}.$$

La primera igualdad es válida por la definición de $f + g$ y la segunda por la relación entre matriz y función, la última es válida por la regla general $AC + BC = (A + B)C$ para la multiplicación y suma matriciales.

Para demostrar (ii), escribimos:

$$rf(\mathbf{x}) = r(A\mathbf{x}) = (rA)\mathbf{x},$$

donde la última igualdad se deduce de la regla $B(AC) = (BA)C$ para la multiplicación matricial.

(iii) es simplemente el Teorema 4.5 del Cap. 1.

Ejemplo 2. Sean f y g funciones lineales tales que

$$f\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix}, \quad f\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix}$$
$$g\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}, \quad g\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\-1\end{pmatrix}.$$

Entonces, de acuerdo con el Teorema 4.2 del Cap. 1, las matrices de f y g son

$$\begin{pmatrix}2 & -1\\3 & 0\end{pmatrix} \quad \text{y} \quad \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$$

respectivamente. Por el teorema recientemente demostrado, la matriz de $f + g$ es la suma de las matrices de f y g:

$$\begin{pmatrix}2 & -1\\3 & 0\end{pmatrix} + \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} = \begin{pmatrix}3 & -1\\3 & -1\end{pmatrix}$$

Por tanto, $f + g$ es una función lineal tal que $f\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix}$ y $f\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-1\\-1\end{pmatrix}$.

En forma análoga, la matriz de $f - 2g$ es

$$\begin{pmatrix}2 & -1\\3 & 0\end{pmatrix} - 2\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} = \begin{pmatrix}0 & -1\\3 & 2\end{pmatrix}.$$

Por último, la matriz de $g \circ f$ es el producto

$$\begin{pmatrix}2 & -1\\3 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} = \begin{pmatrix}2 & 1\\3 & 0\end{pmatrix}.$$

La conexión que existe entre combinaciones de matrices y combinaciones de funciones de $\mathfrak{R}^n$ en $\mathfrak{R}^m$ sugiere que las reglas del álgebra de matrices deben ser válidas para las funciones lineales. En efecto, estas leyes son válidas para funciones lineales cualesquiera, tales que las operaciones estén definidas.

5.2 Teorema

Sea r un número y sean f, g y h funciones lineales para las cuales las combinaciones indicadas más adelante están definidas. Entonces,

(1) $(f + g) \circ h = f \circ h + g \circ h$
(2) $f \circ (g + h) = f \circ g + f \circ h$
(3) $(rf) \circ g = r(f \circ g) = f \circ (rg)$
(4) $h \circ (g \circ f) = (h \circ g) \circ f$.

Demostración. Todas estas fórmulas se demuestran aplicando la función de cada miembro de la igualdad a un vector arbitrario $\mathbf{x}$ del dominio y escribiendo el significado de cada expresión. Demostraremos solo la Ec. (4), dejando las otras como ejercicios. Para ello observamos que, por definición,

$$h \circ (g \circ f)(\mathbf{x}) = h(g \circ f(\mathbf{x}))$$
$$= h(g(f(\mathbf{x}))).$$

Análogamente,

$$(h \circ g) \circ f(\mathbf{x}) = h \circ g(f(\mathbf{x}))$$
$$= h(g(f(\mathbf{x}))).$$

El resultado de estos dos cálculos es el mismo, de modo que las fórmulas de las cuales partimos son iguales. Como $\mathbf{x}$ es arbitrario, $h \circ (g \circ f) = (h \circ g) \circ f$.

Ejemplo 3. Sea $C^{(\infty)}(\mathfrak{R})$ el espacio vectorial de las funciones reales y de la variable x con derivadas de todo orden. Si D denota la derivación con respecto de x, entonces

$$Dy = y',$$

y, evidentemente, D es una función lineal con dominio $C^{(\infty)}(\mathfrak{R})$. Como el significado será siempre claro, omitiremos el círculo al escribir composiciones,

de modo que DD significará $D \circ D$. Las compuestas $D^2 = DD$, $D^3 = DDD$, etc., tienen el significado claro, por ejemplo,

$$D^2y = y'', \quad D^3y = y'''.$$

Además, podemos definir

$$D^0y = y$$

lo cual es conveniente a veces. Combinando potencias de D por adición y multiplicación numérica se obtienen fórmulas tales como

$$D^2 - D - 2, \quad (D + 1)(D - 2),$$

y, aplicando las reglas (1) y (2) del Teorema 4.2, obtenemos

$$(D + 1)(D - 2) = D^2 - D - 2.$$

Si aplicamos ambos miembros de esta ecuación a la función y, obtenemos en ambos casos

$$y'' - y' - 2y.$$

En la próxima sección estudiaremos en más detalle operadores diferenciales de este tipo.

La operación de encontrar la inversa A^{-1} de una matriz invertible A es especialmente importante. La idea de la inversa de una matriz está relacionada con la idea de inversa de una función. Una función f tiene una **función inversa** denotada por f^{-1} si

$$f^{-1} \circ f(\mathbf{x}) = \mathbf{x}$$

para todo $\mathbf{x}$ en el dominio de f. Si aplicamos f a ambos lados de esta ecuación tenemos que $f \circ f^{-1}(f(\mathbf{x})) = f(\mathbf{x})$, de modo que también tenemos que

$$f \circ f^{-1}(\mathbf{y}) = \mathbf{y},$$

para todo $\mathbf{y} = f(\mathbf{x})$ en la imagen de f. Dejamos como ejercicio la demostración de

5.3 Si f es lineal, entonces f^{-1} también lo es.

Como las técnicas para calcular matrices inversas han sido desarrolladas en varias etapas distintas, resumiremos aquí lo que hemos demostrado hasta ahora. Al final del Cap. 1 demostramos

5.4 Si A es una matriz ortogonal, entonces $A^{-1} = A^t$, donde A es la matriz transpuesta de A.

En el Teorema 8.3 del Cap. 1 se demostró que

5.5 Si $\det a \neq 0$, entonces A es invertible y $A^{-1} = (1/\det A)\tilde{A}^t$, donde $\tilde{A}$ es la matriz de cofactores de elementos de A.

Finalmente, el método de reducción por filas de la primera sección de este

capítulo es un método eficiente para encontrar la inversa de una matriz invertible, es decir, encontrar A^{-1} tal que

$$A^{-1}A = AA^{-1} = I,$$

donde I es la matriz identidad.

La relación entre la función lineal inversa y la matriz inversa no es tan simple como en el caso de las otras operaciones entre funciones y matrices. El ejemplo siguiente muestra por qué.

Ejemplo 4. Sea A una matriz n-por-b invertible. Entonces la función lineal f de $\mathcal{R}^n$ en $\mathcal{R}^n$, definida por

$$f(\mathbf{x}) = A\mathbf{x},$$

tiene siempre una función inversa a f^{-1} dada por

$$f^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}.$$

En efecto,

$$f^{-1} \circ f(\mathbf{x}) = A^{-1}A\mathbf{x} = \mathbf{x},$$

para todo $\mathbf{x}$ en $\mathcal{R}^n$. Sin embargo, si consideramos la función lineal g de $\mathcal{R}$ en $\mathcal{R}^2$ definida por

$$g(x) = \begin{pmatrix} x \\ x \end{pmatrix}.$$

es fácil comprobar que g es lineal y que tiene una función inversa definida en el subespacio $\mathcal{L}$ de $\mathcal{R}^2$ formado por los puntos de $\mathcal{R}^2$ que tienen coordenadas iguales.

Definimos

$$g^{-1}\begin{pmatrix} x \\ x \end{pmatrix} = x, \quad \text{para} \quad \begin{pmatrix} x \\ x \end{pmatrix} \text{ en } \mathcal{L}.$$

Pero la matriz de g es

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

que no es invertible, ya que ni siquiera es cuadrada. La dificultad está en que g^{-1} no está definida en todo $\mathcal{R}^2$, sino solo en el subespacio $\mathcal{L}$.

En el ejemplo anterior, la característica esencial de la función g es la dimensión de su imagen. Esta idea es lo suficientemente importante como para tener un nombre, el **rango** de una función lineal es la dimensión de la imagen de f. Así, por ejemplo, la función lineal

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \begin{pmatrix} 2x \\ 3y \end{pmatrix}$$

tiene rango igual a 2, ya que su imagen es todo $\mathfrak{R}^2$. La función

$$g\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\0\end{pmatrix}$$

tiene rango 1, ya que su imagen es el eje x en $\mathfrak{R}^2$. Para una función lineal f de $\mathfrak{R}^n$ en $\mathfrak{R}^m$ con matriz A, se tiene el siguiente útil criterio.

5.6 Teorema

Sea f una función lineal con matriz A. Entonces el rango de f es igual al número de columnas linealmente independientes de A.

Demostración. Las columnas de A generan la imagen de f, ya que

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + \ldots + x_n f(\mathbf{e}_n),$$

donde $\mathbf{x} = (x_1, \ldots, x_n)$ y los vectores columna $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ son simplemente las columnas de A, de modo que estas columnas generan la imagen de f. Por tanto, el número de columnas independientes es la dimensión de la imagen de f.

Por el Teorema 5.6, podemos hablar del **rango de una matriz** A y definirlo ya sea como el número de columnas linealmente independientes o como el rango de la función $f(\mathbf{x}) = A\mathbf{x}$.

Ejemplo 5. Para calcular el rango de la matriz

$$\begin{pmatrix}1 & 0 & 7\\0 & 1 & 5\\2 & -1 & 9\end{pmatrix},$$

aplicamos operaciones elementales a las filas para reducir la matriz a su forma escalonada. En cada etapa, la nueva matriz es la matriz de una función lineal distinta, pero como las operaciones elementales son invertibles, la dimensión de la imagen no cambia. Tenemos que, sumando la segunda línea a la tercera, obtenemos

$$\begin{pmatrix}1 & 0 & 7\\0 & 1 & 5\\2 & 0 & 14\end{pmatrix}.$$

Entonces, restando 2 veces la primera fila de la tercera, obtenemos

$$\begin{pmatrix}1 & 0 & 7\\0 & 1 & 5\\0 & 0 & 0\end{pmatrix}.$$

No es posible hacer nuevas reducciones de modo que, evidentemente, la matriz tiene dos columnas linealmente independientes. Por tanto, la matriz dada tiene rango 2 y la función lineal de $\mathfrak{R}^3$ en $\mathfrak{R}^3$ asociada a la matriz tiene una imagen 2-dimensional.

Por último, completamos la discusión de la relación entre la función inversa y la matriz inversa en $\mathfrak{R}^n$. Si f es una función, entonces se dice que f es **inyectiva** si a cada elemento de la imagen corresponde exactamente un elemento del dominio de f. Es claro que f es inyectiva si, y solo si, f^{-1} existe. Para funciones *lineales* tenemos el siguiente

5.7 Teorema

Una función lineal f es inyectiva si, y solo si, $\mathbf{x} = 0$ cada vez que $f(\mathbf{x}) = 0$.

Demostración. Primero supongamos que $f(\mathbf{x}) = 0$ implica que $\mathbf{x} = 0$. Si $f(\mathbf{x}_1) = f(\mathbf{x}_2)$, entonces, por la linealidad de f, $f(\mathbf{x}_1 - \mathbf{x}_2) = 0$. Pero entonces, por hipótesis, $\mathbf{x}_1 - \mathbf{x}_2 = 0$, de modo que $\mathbf{x}_1 = \mathbf{x}_2$. Por tanto, f es inyectiva. Conversamente, supongamos que $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ implica que $\mathbf{x}_1 = \mathbf{x}_2$. Entonces como para una función lineal $f(0) = 0$, la ecuación $f(\mathbf{x}) = 0$ puede ser escrita en la forma $f(\mathbf{x}) = f(0)$. Pero entonces, por hipótesis, $\mathbf{x} = 0$.

Del Teorema 5.7 se deduce que si f es una función lineal inyectiva de $\mathfrak{R}^n$ en $\mathfrak{R}^m$, entonces el rango de f es n. La razón para esto es que el núcleo de f tiene dimensión 0 (es decir, la dimensión del origen), de modo que, por el Teorema 4.7, rango $(f) + 0 = n$. Esta idea se usa en la demostración del teorema siguiente. Para funciones lineales de $\mathfrak{R}^n$ en $\mathfrak{R}^n$, es decir, para funciones representables por matrices cuadradas, tenemos

5.8 Teorema

Sea $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^n$ una función lineal con matriz cuadrada A. Entonces f tiene inversa si, y solo si, A es invertible.

Demostración. Hemos visto al comienzo del Ejemplo 4 que si A^{-1} existe, entonces f^{-1} también existe. Conversamente, si f tiene una inversa f^{-1}, entonces f es inyectiva. Por el Teorema 5.7, el núcleo de f tiene dimensión 0, de modo que la imagen de f es todo $\mathfrak{R}^n$. Pero esto significa que el dominio de f^{-1} es todo $\mathfrak{R}^n$. De aquí sigue que f^{-1} tiene una matriz cuadrada B tal que $f^{-1}(\mathbf{x}) = B\mathbf{x}$ para todo $\mathbf{x}$ en $\mathfrak{R}^n$. Entonces las ecuaciones

$$BA\mathbf{x} = f^{-1} \circ f(\mathbf{x}) = \mathbf{x}$$
$$AB\mathbf{x} = f \circ f^{-1}(\mathbf{x}) = \mathbf{x}$$

son válidas. Por tanto, $AB = BA$, de modo que A es invertible y $A^{-1} = B$.

EJERCICIOS

1. Suponga que f y g son funciones lineales tales que

$$f\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}, \quad f\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix},$$

$$g\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix}, \quad g\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-2\\3\end{pmatrix}.$$

Encuentre las matrices de las funciones siguientes

(a) f.
(b) g.
(c) $f + g$.
(d) $2f - g$.
(e) $g \circ f$.
(f) $f \circ (f + g)$.

2. Si

$$A = \begin{pmatrix}1 & 1\\0 & 2\\1 & 3\end{pmatrix}, \quad B = \begin{pmatrix}-1 & 0\\2 & 1\\3 & 1\end{pmatrix}, \quad C = \begin{pmatrix}2 & 1\\1 & 3\end{pmatrix},$$

calcule $(A + B)C$ y verifique que es igual a $AC + BC$.

3. (a) Demuestre las reglas (1), (2) y (3) del Teorema 5.2.
(b) Demuestre que las combinaciones lineales y las compuestas de funciones lineales son lineales.

4. Encuentre las inversas de cada una de las matrices dadas más adelante, si ellas existen.

(a) $\begin{pmatrix}1 & 3\\-1 & 2\end{pmatrix}$.

(b) $\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}$.

(c) $\begin{pmatrix}1 & 0 & 2\\0 & 1 & 2\\-3 & 3 & 0\end{pmatrix}$.

(d) $\begin{pmatrix}2 & -1 & 3\\1 & 2 & 1\\2 & -2 & 1\end{pmatrix}$.

5. Demuestre que si f es una función lineal que tiene una inversa f^{-1}, entonces f^{-1} también es lineal.

6. Sea f una función lineal de $\mathcal{R}^n$ en $\mathcal{R}^m$.
(a) Demuestre que si $n > m$, entonces la dimensión del núcleo de f es positiva. Demuestre entonces que f no puede tener una función inversa.
(b) Suponga que $n = m$ y que f tiene una función inversa f^{-1}. Demuestre que si A es la matriz de f, entonces A es invertible y A^{-1} es la matriz de f^{-1}.
(c) Demuestre que si $n < m$, entonces aunque f^{-1} exista, la matriz de f no es invertible.

7. (a) Demuestre que $(D + a)(D + b) = (D + b)(D + a)$, donde a y b son constantes.
(b) Defina $D + f(x)$ por

$$(D + f(x))y = y' + f(x)y.$$

Demuestre que

$$(D + 1)(D + x) \neq (D + x)(D + 1)$$

aplicando ambos miembros a una función dos veces diferenciable $y(x)$.

8. Encuentre el rango de f, si f tiene la matriz

(a) $\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ (b) $\begin{pmatrix}1 & 2\\2 & 4\end{pmatrix}$

(c) $\begin{pmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{pmatrix}$ (d) $\begin{pmatrix}2 & 1 & 3\\2 & -2 & 2\\4 & -1 & 5\end{pmatrix}$.

9. Sean $\mathcal{U}$ y $\mathcal{W}$ dos espacios vectoriales de dimensión finita y sea f una función lineal con dominio $\mathcal{U}$ y cuya imagen es $\mathcal{W}$. (Por tanto, dim $(\mathcal{W})$ = rango (f).)
(a) Demuestre que $\mathcal{U}$ contiene un subespacio $\mathcal{S}$ tal que f, cuando se la restringe a $\mathcal{S}$, es inyectiva de $\mathcal{S}$ en $\mathcal{W}$. [*Indicación:* Sea $\mathcal{N}$ el núcleo de f con la base $\mathbf{n}_1, \ldots, \mathbf{n}_k$. Extienda esta base a una base para $\mathcal{W}$ agregando vectores $\mathbf{s}_1, \ldots, \mathbf{s}_l$.]
(b) Demuestre que dim $(\mathcal{S})$ = rango (f), donde $\mathcal{S}$ es el subespacio de la parte (a).

10. Si f es una función e $\mathbf{y}$ es un elemento de su imagen, entonces el subconjunto S del dominio de f formado por todos los $\mathbf{x}$ para los cuales $f(\mathbf{x}) = \mathbf{y}$ se llama la **imagen inversa** o **preimagen** de $\mathbf{y}$. Demuestre que si f es lineal entonces la imagen inversa de un vector fijo es un subespacio del dominio de f o un subespacio transladado por un vector fijo, es decir, un subespacio afín.

11. ¿Cuál es la manera más simple de calcular el rango de la matriz diagonal **diag** $(a_1, \ldots, a_n)$?

12. Demuestre que si A es una matriz m-por-n, entonces el rango de A es $n - k$, donde k es la dimensión del conjunto de soluciones de la ecuación $A\mathbf{x} = 0$.

6. Operadores diferenciales

En esta sección consideramos algunas de las ideas sobre espacios vectoriales que surgen en el estudio de las ecuaciones diferenciales. Las ecuaciones que consideraremos son del tipo

$$y' - 2y = 0 \tag{1}$$

$$y' - 3y = e^x \tag{2}$$

$$y'' + 3y' + y = \operatorname{sen} x, \tag{3}$$

donde los ápices denotan derivación con respecto de x. Las ecuaciones de este tipo son importantes, ya que ellas expresan las relaciones entre el valor de y, su primera derivada (velocidad), y su segunda derivada (aceleración). El problema planteado por cada una de estas ecuaciones es el de encontrar todas las funciones $y(x)$ tales que, remplazando y por $y(x)$ se tiene una identidad. Antes de empezar un estudio sistematico del problema, consideraremos algunos ejemplos que serán útiles más adelante.

Ejemplo 1. La ecuación diferencial $y' - ry = 0$, donde r es una constante, puede ser escrita en la forma

$$y' = ry \tag{4}$$

y, por tanto, expresa que la razón de variación de y es proporcional al valor de y para todo valor de la variable x. A veces se supone que el crecimiento de una población satisface una ley de este tipo. Para encontrar soluciones usamos el hecho de que si $y = e^{rx}$, entonces $y' = re^{rx}$. De esto se deduce que la Ec. (4) se satisface tomando $y = e^{rx}$. Más generalmente, si c es una constante arbitraria, entonces la Ec. (4) es satisfecha por

$$y = ce^{rx}, \tag{5}$$

ya que c se cancelará. Es más, la Ec. (5) es la solución más general de (4). En efecto, nótese que (5) puede ser escrita en la forma

$$e^{-rx}y = c.$$

Derivando con respecto de x se tiene

$$(e^{-rx}y)' = 0$$

o, usando la regla del producto para derivadas,

$$e^{-rx}y' - re^{-rx}y = e^{-rx}(y' - ry) = 0.$$

Dividiendo por e^{-rx} se tiene $y' - ry = 0$ que es lo mismo que la Ec. (4). Pero ahora podemos invertir estos pasos suponiendo que y es *alguna* solución. Empezamos por

$$y' - ry = 0$$

y luego multiplicamos por e^{-rx} para obtener

$$e^{-rx}y' - re^{-rx}y = 0.$$

Por la regla del producto, esta ecuación es

$$(e^{-rx}y)' = 0.$$

Integrando ambos miembros con respecto de x, tenemos

$$e^{-rx}y = c,$$

donde c es una constante de integración. Multiplicando ambos miembros por e^{rx} demuestre que y debe tener la forma

$$y = ce^{rx}.$$

Por tanto, hemos demostrado que ce^{rx} es la solución más general de $y' = ry$ en el sentido de que toda solución particular se puede obtener especificando un valor de c.

El método usado en el ejemplo anterior consiste en multiplicar la expresión $y' + ay$ por e^{ax} reconociendo luego que el resultado es la derivada $(e^{rx}y)' = e^{ax}y' + ae^{ax}y$.

En el futuro usaremos repetidamente este **multiplicador exponencial** e^{ax}.

Ejemplo 2. Para resolver la ecuación diferencial

$$y' - 3y = e^x,$$

multiplicamos por e^{-3x}, obteniendo

$$e^{-3x}y' - 3e^{-3x}y = e^{-2x}.$$

Esto es lo mismo que

$$(e^{-3x}y)' = e^{-2x}.$$

Ahora integramos ambos miembros con respecto de x, obteniendo

$$e^{-3x}y = -\tfrac{1}{2}e^{-2x} + c,$$

donde c es una constante de integración. Entonces, multiplicando por e^{3x}, tenemos que

$$y = -\tfrac{1}{2}e^x + ce^{3x}$$

es la solución más general. Es fácil verificar directamente que hemos encontrado *algunas* soluciones, una para cada valor de c. Además, hemos demostrado que *toda* solución debe ser de la forma $-\frac{1}{2}e^x + ce^{3x}$.

Antes de considerar ejemplos más complicados, es conveniente describir ciertas notaciones usadas a menudo en la resolución de ecuaciones diferenciales. D significará derivación con respecto de una variable determinada que designamos por x. Entonces interpretamos $D + 2$, $D^2 - 1$ y otras expresiones similares como funciones lineales que actúan sobre funciones y con suficientes derivadas. Por ejemplo:

$$\begin{aligned}(D + 2)y &= Dy + 2y \\ &= y' + 2y \\ (D^2 - 1)y &= D^2y - y \\ &= D(Dy) - y = y'' - y.\end{aligned}$$

Una observación importante es que D actúa como una *función lineal* en y. Para evitar posibles confusiones con el hecho de que y es una función de x (no

necesariamente lineal) a veces se usa el término **operador lineal**. Para comprobar que D actúa linealmente basta recordar las propiedades usuales de la derivación:

$$D(y_1 + y_2) = Dy_1 + Dy_2$$
$$D(ky) = kDy, \qquad k \text{ constante}$$

Estas dos ecuaciones expresan la linealidad de D. Como la composición de funciones lineales es lineal, los operadores D^2, D^3 y, en general D^n, son lineales. Como la multiplicación numérica es una operación lineal y la suma de operaciones lineales es lineal, el operador $D + a$ es lineal para toda constante a. Resumiendo, tenemos que las expresiones de la forma

$$D^2 + a, \qquad D^2 + aD + b, \qquad (D + s)(D + t)$$

son todos operadores lineales. Las interpretaciones de estos operadores son

$$(D^2 + a)y = y'' + ay$$
$$(D^2 + aD + b)y = y'' + ay' + by$$
$$\begin{aligned}(D + s)(D + t)y &= (D + s)(y' + ty)\\ &= D(y' + ty) + s(y' + ty)\\ &= y'' + ty' + sy' + sty\\ &= y'' + (t + s)y' + sty\\ &= \big(D^2 + (s + t)D + st\big)y.\end{aligned}$$

El último ejemplo muestra que si s y t son constantes,

$$(D + s)(D + t) = D^2 + (s + t)D + st,$$

y que $(D + t)(D + s) = D^2 + (s + t)D + st$. Por tanto, los operadores de la forma $D + a$ pueden ser multiplicados como polinomios en D si a es una constante. En la dirección opuesta podemos tratar de factorizar un operador, por ejemplo, $D^2 - 1$. Es claro que

$$\begin{aligned}D^2 - 1 &= (D - 1)(D + 1)\\ &= (D + 1)(D - 1).\end{aligned}$$

Volviendo a considerar ecuaciones diferenciales, supongamos que tenemos una de la forma

$$y'' + ay' + by = 0;$$

La Ec. (3) del comienzo de la sección es de este tipo con $a = 3$, $b = 1$. Si escribimos esta ecuación usando operadores diferenciales tenemos

$$(D^2 + aD + b)y = 0.$$

Nuestro método de solución consistirá en tratar de factorizar el operador en factores de la forma $(D + s)$ y $(D + t)$ para después aplicar repetidamente el método del multiplicador exponencial que usáramos en los Ejemplos 1 y 2.

Ejemplo 3. Supongamos que queremos encontrar todas las funciones $y = y(x)$ que satisfacen

$$y'' + 5y' + 6y = 0.$$

Usando operadores, la ecuación es

$$(D^2 + 5D + 6)y = 0.$$

Ahora factorizamos el operador en la forma

$$(D^2 + 5D + 6) = (D + 3)(D + 2);$$

Por tanto, vemos que debemos resolver

$$(D + 3)(D + 2)y = 0.$$

Para encontrar todas las soluciones supondremos que y es *una* solución. Haciendo

$$(D + 2)y = u$$

sustituimos u en la ecuación anterior, obteniendo

$$(D + 3)u = 0.$$

Pero esta ecuación puede ser resuelta en u multiplicando por e^{3x}. Obtenemos

$$e^{3x}Du + 3e^{3x}u = 0$$

es decir,

$$D(e^{3x}u) = 0.$$

Por tanto

$$e^{3x}u = c_1,$$

para cierta constante c_1, de modo que

$$u = c_1e^{-3x}.$$

Recordemos que hemos tomado $(D + 2)y = u$. Por tanto,

$$(D + 2)y = c_1e^{-3x}.$$

Multiplicando esta ecuación por e^{2x}, tenemos

$$e^{2x}Dy + 2e^{2x}y = c_1e^{-x}$$

es decir,

$$D(e^{2x}y) = c_1e^{-x}.$$

Integrando con respecto a x, tenemos

$$e^{2x}y = -c_1e^{-x} + c_2$$

es decir,

$$y = -c_1e^{-3x} + c_2e^{-2x}.$$

Como las constantes c_1 y c_2 son arbitrarias, podemos cambiar el signo de la primera y entonces

$$y = c_1e^{-3x} + c_2e^{-2x}$$

es la solución más general.

El multiplicador exponencial usado en los ejemplos anteriores se encuentra como sigue: se multiplica $(D + a)y$ por e^{ax} para obtener $D(e^{ax}y)$, es decir,

$$e^{ax}(D + a)y = D(e^{ax}y).$$

La aplicación repetida de esta regla conduce al siguiente resultado general.

6.1 Teorema

La solución más general de la ecuación diferencial

$$(D - r_1)(D - r_2) \ldots (D - r_n)y = 0,$$

donde $r_1, r_2, \ldots, r_n$ son *todos diferentes*, es de la forma

$$y = c_1e^{r_1x} + c_2e^{r_2x} + \ldots + c_ne^{r_nx}.$$

Si algunos de los r son iguales, digamos $r_1 = r_2 = r_3 = \ldots = r_k$, remplazamos $e^{r_2x}, e^{r_3x}, \ldots, e^{r_kx}$ por $xe^{r_1x}, x^2e^{r_2x}, \ldots, x^{k-1}e^{r_1x}$, respectivamente, obteniendo así la solución general.

Demostración. Por simplicidad, empezaremos por el caso $n = 2$. Dada

$$(D - r_1)(D - r_2)y = 0,$$

hacemos

$$(D - r_2)y = u.$$

Entonces la ecuación

$$(D - r_1)u = 0$$

se resuelve multiplicando ambos miembros por e^{-r_1x}. Obtenemos

$$e^{-r_1x}Du - r_1e^{-r_1x}u = 0$$

es decir,

$$D(e^{-r_1x}u) = 0.$$

Entonces, integrando con respecto a x obtenemos

$$e^{-r_1x}u = c_1$$

es decir,

$$u = c_1e^{r_1x}.$$

Ahora bien,

$$(D - r_2)y = c_1e^{r_1x}.$$

Multiplicamos por e^{-r_2x} y obtenemos

$$e^{-r_2x}Dy - r_2e^{-r_2x}y = c_1e^{(r_1-r_2)x}.$$

Integrando ambos miembros de

$$D(e^{-r_2x}y) = c_1e^{(r_1-r_2)x} \tag{6}$$

se tiene que

$$e^{-r_2x}y = \frac{c_1}{r_1 - r_2}e^{(r_1-r_2)x} + c_2$$

es decir,

$$y = \frac{c_1}{r_1 - r_2}e^{r_1x} + c_2e^{r_2x}.$$

Hemos supuesto que $r_1 \neq r_2$, de modo que la última integración es correcta. Para mayor elegancia, remplazamos la constante arbitraria c_1 por $(r_1 - r_2)c_1$ que es tan arbitraria como c_1.

Si $r_1 = r_2$, entonces la última integración no es correcta, ya que en este caso $r_1 - r_2 = 0$ y la Ec. (6) toma la forma

$$D(e^{-r_2x}y) = c_1.$$

Si integramos, obtenemos

$$e^{-r_2x}y = c_1x + c_2$$

es decir,

$$y = c_1xe^{r_2x} + c_2e^{r_2x},$$

como lo afirma el teorema.

Más generalmente, para resolver

$$(D - r_1)(D - r_2) \ldots (D - r_n)y = 0,$$

hacemos

$$(D - r_2) \ldots (D - r_n)y = u_1, \tag{7}$$

de modo que la ecuación anterior toma la forma

$$(D - r_1)u_1 = 0.$$

Sustituyendo en (7) la solución general para u_1 obtenemos una nueva ecuación que la separamos escribiendo

$$(D - r_3) \ldots (D - r_n)y = u_2.$$

Entonces la Ec. (7) se transforma en

$$(D - r_2)u_2 = u_1,$$

en la cual hay que encontrar la solución u_2. Continuando de esta manera, finalmente tenemos que encontrar la solución y de la ecuación

$$(D - r_n)y = u_{n-1}$$

La solución de cada una de estas ecuaciones se reduce a la solución de ecuaciones de la forma

$$(D - r)y = cx^ke^{sx}, \qquad s \neq r$$

o

$$(D - r)y = cx^ke^{rx}.$$

En el primer caso multiplicamos por e^{-rx}. Entonces, integrando por partes, se obtiene una combinación lineal de potencias de x multiplicadas por e^{sx}. En el segundo caso, después de multiplicar por e^{-rx}, solo tenemos que integrar x^k en la derecha. En ambos casos obtenemos soluciones de la forma indicada en el teorema.

La clave de la aplicación de este teorema es la factorización de un operador diferencial en factores de la forma $(D - r)$. Considerado como un polinomio en D, la expresión

$$D^n + a_{n-1}D^{n-1} + \ldots + a_1 D + a_0 \tag{8}$$

se llama el **polinomio característico** de la ecuación diferencial asociada,

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0.$$

Nótese que para obtener el polinomio a partir de la ecuación, remplazamos $y^{(k)}$ por D^k y que el término $a_0 y$ corresponde a $a_0 D^0 = a_0$. Para encontrar la factorización del polinomio es necesario conocer sus raíces, ya que si el polinomio (7) tiene las raíces $r_1, \ldots, r_n$, entonces se lo puede factorizar en la forma

$$(D - r_1)(D - r_2) \ldots (D - r_n).$$

En general, es imposible encontrar exactamente las raíces de un polinomio. Sin embargo, si $n = 2$ (o aun 3 o 4) hay fórmulas bien conocidas para encontrar las raíces en términos de los coeficientes. Por simplicidad, hemos supuesto que el primer coeficiente del polinomio característico es 1, de no ser así, basta con dividir todos los términos para obtener uno con primer coeficiente 1.

Ejemplo 4. Supongamos que queremos resolver

$$y''' - 4y'' + 4y' = 0.$$

El polinomio característico es

$$D^3 - 4D^2 + 4D,$$

que puede ser escrito en la forma

$$D(D^2 - 4D + 4) = D(D - 2)^2$$

Las raíces son 0 y 2, donde 2 es una raíz repetida. Por tanto, la solución general es una combinación lineal de e^{0x}, e^{2x} y xe^{2x}, de modo que la solución general es

$$y = c_1 + c_2 e^{2x} + c_3 x e^{2x}$$

De aquí en adelante describiremos la teoría en términos de espacios vectoriales. Nos hemos referido a la linealidad de los operadores $(D - r)$ sin especificar el espacio vectorial en que ellos operan. Hay varias posibilidades que dependen de los requisitos de los problemas particulares. Una posibilidad es considerar el espacio vectorial $C^{(n)}$ de las funciones reales de x con derivada n-ésima continua. Si L es un operador diferencial de orden n, es decir, un operador que contiene derivaciones de orden a lo más n, entonces L es lineal de $C^{(n)}$ en $C^{(0)}$, el espacio de las funciones continuas. Otra posibilidad es considerar que L actúa en el espacio vectorial $C^{(\infty)}$ de las funciones con derivadas de todo orden. Para nuestros fines, esta última elección será más natural. Nótese que las funciones e^{rx} y $x^k e^{rx}$ del Teorema 6.1 pertenecen a $C^{(\infty)}$ y, por tanto, también a $C^{(n)}$ para todo n. Además, el conjunto $\mathfrak{N}$ de las soluciones de

$$(D - r_1) \ldots (D - r_n)y = 0.$$

es un subespacio vectorial de $C^{(n)}$, ya que $\mathfrak{N}$ es el núcleo del operador lineal

$$L = (D - r_1) \ldots (D - r_n)$$

que actúa en $C^{(n)}$.

Como $\mathfrak{N}$ es generado por n funciones distintas de la forma e^{rx} o $x^j e^{rx}$, del Teorema 6.1 se deduce inmediatamente que $\mathfrak{N}$ tiene dimensión no mayor que el orden n de L. (No se excluye la posibilidad que r sea un número complejo. El significado de este caso será explicado en la próxima sección.) De hecho, se puede demostrar el teorema siguiente, cuya demostración consiste simplemente en demostrar que las soluciones básicas son linealmente independientes, lo cual requiere algún trabajo.

6.2 Teorema

Sea $\mathfrak{N}$ el subespacio de $C^{(n)}$, formado por todas las soluciones de la ecuación diferencial lineal

$$(D - r_1) \ldots (D - r_n)y = 0$$

Entonces $\mathfrak{N}$ tiene dimensión n.

Demostración. Ya hemos notado que la dimensión de $\mathfrak{N}$ no es mayor que n, puesto que es generado por un conjunto con n elementos. Podemos demostrar que la dimensión es n revisando la forma en que funciona el método del multiplicador exponencial. Debemos resolver una sucesión de ecuaciones diferenciales

$$(D - r_k)y = u_{k-1}$$

donde u_{k-1} tiene la forma general

$$u_{k-1}(x) = c_1 x^{l_1} e^{r_1 x} + \ldots + c_{k-1} x^{l_{k-1}} e^{r_{k-1} x}.$$

Aplicamos el factor $e^{-r_k x}$ para obtener, como de costumbre,

$$D(e^{-r_k x}y) = e^{-r_k x}u_{k-1}. \tag{9}$$

Si suponemos, por inducción, que el conjunto de las soluciones u_{k-1} tiene dimensión $k - 1$, entonces lo mismo es verdad para el conjunto de funciones $e^{-r_k x}u_{k-1}$. (¿Por qué?) Por tanto, consideramos que el operador lineal D tiene como dominio las funciones $e^{-r_k x}y$ y su espacio imagen es de dimensión $k - 1$. Pero el núcleo de D tiene dimensión 1, puesto que las soluciones de

$$D(e^{-r_k x}y) = 0$$

son de la forma

$$e^{-r_k x}y = c,$$

lo cual es evidentemente un subespacio de dimensión 1 del dominio D. Como

$$\dim(\text{dominio}) = \dim(\text{núcleo}) + \dim(\text{imagen})$$

tenemos que dim (dominio) $= 1 + (k - 1) = k$. Por tanto, el conjunto de soluciones $e^{-r_k x}y$ de la Ec. (9) tiene dimensión k. Por tanto, el conjunto de funciones de la forma

$$y = e^{r_k x}(e^{-r_k x}y)$$

también tiene dimensión k. Como hemos demostrado esto para $k = 1, 2, \ldots, n$, hemos demostrado que el conjunto $\mathfrak{N}$ de todas las soluciones de la ecuación diferencial tiene dimensión n.

Concluimos esta sección considerando ecuaciones diferenciales de la forma

$$L(y) = f,$$

donde f es una función continua de x que se supone dada y L tiene la forma

$$L = (D - r_1)(D - r_2)\ldots(D - r_n).$$

Ya hemos considerado el caso en que f es idénticamente cero.

Ejemplo 5. Dado

$$y'' + 2y' + y = e^{3x},$$

escribimos el polinomio característico $D^2 + 2D + 1$ y lo factorizamos, con lo cual la ecuación toma la forma

$$(D + 1)^2 y = e^{3x}.$$

Tomando $(D + 1)y = u$, tratamos de resolver

$$(D + 1)u = e^{3x}.$$

Multiplicando por e^x, obtenemos

$$e^x Du + e^x u = e^{4x}$$

es decir,

$$D(e^x u) = e^{4x}.$$

Entonces, integrando, se tiene

$$e^x u = \tfrac{1}{4}e^{4x} + c_1$$

es decir,

$$u = \tfrac{1}{4}e^{3x} + c_1 e^{-x}.$$

Como $(D + 1)y = u$, tenemos que

$$(D + 1)y = \tfrac{1}{4}e^{3x} + c_1 e^{-x}.$$

Multiplicando nuevamente por e^x, obtenemos

$$e^x Dy + e^x y = \tfrac{1}{4}e^{4x} + c_1$$

es decir,

$$D(e^x y) = \tfrac{1}{4}e^{4x} + c_1.$$

Entonces,

$$e^x y = \tfrac{1}{16}e^{4x} + c_1 x + c_2$$

o bien

$$y = \tfrac{1}{16}e^{3x} + c_1 x e^{-x} + c_2 e^{-x}.$$

En el ejemplo anterior, la solución es la suma de dos partes, y_h e y_p:

$$y_h = c_1 x e^{-x} + c_2 e^{-x},$$
$$y_p = \tfrac{1}{16}e^{3x}.$$

La función y_h se llama la **parte homogénea** de la solución, ya que es la solución de la **ecuación homogénea** asociada

$$L(y) = 0$$

La función y_p se llama **solución particular de**

$$L(y) = f$$

puesto que no es la solución más general. De hecho, obtenemos y_p haciendo $c_1 = c_2 = 0$ en la solución general. Esta división de la solución en dos partes es un ejemplo de un hecho general acerca de las funciones lineales. Este hecho es sencillo y se utiliza en la solución sistemas lineales algebraicos. Como este principio es importante, lo damos a continuación.

6.3 Teorema

Sea L una función lineal. Sea f un elemento de la imagen de L y sea y_p un elemento del dominio de L tal que $L(y_p) = f$. Entonces todas las soluciones y de

$$L(y) = f$$

pueden ser expresadas como una suma

$$y = y_h + y_p,$$

donde y_h es un elemento del núcleo $\mathfrak{N}$ de L.

Demostración. Supongamos que $L(y) = f$ y que también $L(y_p) = f$. Entonces, como L es lineal,

$$L(y - y_p) = L(y) = L(y_p)$$
$$= f - f = 0.$$

Por tanto, $y - y_p$ está en el núcleo $\mathfrak{N}$ de L, es decir, $y - y_p = y_h$ para algún y_h en $\mathfrak{N}$. Pero entonces $y = y_h + y_p$, como se quería demostrar.

El método del Ejemplo 5 puede ser usado siempre para encontrar la solución más general de la ecuación $L(y) = f$ de la forma

$$(D - r_1)\ldots(D - r_n)y = f.$$

Sin embargo, en algunos ejemplos los cálculos pueden ser acortados conside-

rablemente recurriendo al Teorema 6.1, que dice que la solución y_h de la parte homogénea puede ser encontrada una vez que se conocen las raíces del polinomio característico. Entonces el Teorema 6.3 dice que si encontramos la solución general y_h de la parte homogénea (usando el Teorema 6.1) y de alguna manera encontramos una solución particular y_p, entonces la solución general de la ecuación es $y_h + y_p$. Podrá ser conveniente usar la linealidad de L para encontrar y_p en el caso en que el término de la derecha, f, es una suma de dos o más términos. Si queremos resolver

$$L(y) = a_1 f_1 + a_2 f_2 \tag{10}$$

y podemos encontrar soluciones y_1 e y_2, tales que

$$L(y_1) = f_1, \qquad L(y_2) = f_2,$$

entonces, como L es lineal, la función

$$y = a_1 y_1 + a_2 y_2$$

es una solución de la Ec. (10). En este contexto, la propiedad de linealidad se llama a veces el **principio de superposición,** puesto que la solución se encuentra por superposición (es decir, suma) de las soluciones de más de una ecuación.

Ejemplo 6. En el Ejemplo 5 vimos que la ecuación diferencial

$$(D + 1)^2 y = e^{3x}$$

tiene la solución general

$$\tfrac{1}{16}e^{3x} + c_1 x e^{-x} + c_2 e^{-x}.$$

Si $c_1 = c_2 = 0$, obtenemos la solución particular $y_1 = \frac{1}{16}e^{3x}$.

Si ahora queremos resolver

$$(D + 1)^2 y = e^{3x} + 1, \tag{11}$$

no es necesario empezar desde el principio; basta con encontrar una solución particular de

$$(D + 1)^2 y = 1.$$

Esta ecuación podría ser resuelta usando los multiplicadores exponenciales, pero en este caso la ecuación diferencial es tan sencilla, que se ve de inmediato que $y_2 = 1$ es una solución. Entonces una solución particular de la Ec. (11) es $y_p = \frac{1}{16}e^{3x} + 1$, y la solución general es

$$y = \tfrac{1}{16}e^{3x} + 1 + c_1 x e^{-x} + c_2 e^{-x}.$$

Por otra parte, al resolver

$$(D + 1)^2 y = e^{3x} + e^{-x}$$

necesitamos encontrar una solución particular de

$$(D + 1)^2 y = e^{-x}.$$

Para encontrarla, usamos el método del multiplicador exponencial.

En los Ejemplos 10 y 11 se dan otros métodos de solución usando «coeficientes indeterminados» y «variación de parámetros». A veces, estos métodos son más eficientes para encontrar soluciones particulares.

EJERCICIOS

1. Encuentre un multiplicador exponencial adecuado para cada una de las ecuaciones diferenciales siguientes. resolviendo después la ecuación modificada integrando ambos términos.

   (a) $y' + 2y = 1$. (c) $y' - y = e^x$.

   (b) $2\dfrac{dy}{dx} + y = x$. (d) $y' = \operatorname{sen} x$.

2. Demuestre que si la expresión $y' + p(x)y$ se multiplica por

   $$e^{\int p(x)\,dx},$$

   el producto resultante puede ser escrito en la forma

   $$\frac{d}{dx}\left(e^{\int p(x)\,dx}\, y\right).$$

   [Suponemos que $p(x)$ es una función continua de x y que $\int p(x)\,dx$ es cualquier función con derivada $p(x)$.]

3. Use el resultado del Ejercicio 2 para encontrar un multiplicador exponencial para cada una de las ecuaciones diferenciales siguientes. Resuelva cada ecuación.

   (a) $y' + \left(\dfrac{1}{x}\right)y = 1$. (c) $y' - xy = 0$.

   (b) $\dfrac{dy}{dx} + xy = x$. (d) $y' + y = 0$.

4. Escriba cada una de las siguientes ecuaciones diferenciales usando operadores, por ejemplo, $(D^2 + 2D + 1)y = 0$, y factorice el operador en factores de orden uno. Encuentre la solución general de la ecuación.

   (a) $y'' + 2y' + y = 0$. (e) $(D^2 - 1)y = 1$.

   (b) $2y'' - y = 0$. (f) $y'' - y = e^x$.

   (c) $y''' + 3y'' + y' = 0$. (g) $y'' - y = e^x + 1$.

   (d) $D(D + 3)y = 0$.

5. Dibuje el gráfico de cada una de las funciones de x dadas más abajo. Encuentre una ecuación diferencial de la cual cada una es solución. Las ecuaciones son de la forma $y'' + ay' + by = 0$.

   (a) xe^{-x}. (c) $1 + x$.

   (b) $e^x + e^{-x}$. (d) $2e^{2x} - 3e^{3x}$.

6. Definimos el operador diferencial $D + f(x)$ por $(D + f(x))y = y' + f(x)y$ para toda función y.

   (a) Demuestre que, en general $(D + f(x))(D + g(x)) \neq (D + g(x))(D + f(x))$.
   (b) Demuestre que si f y g son constantes entonces se tiene igualdad en la parte (a).

7. Para $x \neq 0$, la ecuación diferencial

$$x^2y'' + (x^3 + x)y' + (x^2 - 1)y = 0$$

puede ser escrita en forma de operadores como

$$\left(D^2 + \left(\frac{x^2 + 1}{x}\right)D + \left(\frac{x^2 - 1}{x^2}\right)\right)y = 0.$$

(a) Demuestre que esta ecuación puede ser escrita en la forma

$$(D + x)\left(D + \frac{1}{x}\right)y = 0.$$

(b) Resuelva la ecuación de la parte (a) aplicando sucesivamente los multiplicadores exponenciales de la forma dada en el Ejercicio 2.

(c) Demuestre que

$$(D + x)\left(D + \frac{1}{x}\right) \neq \left(D + \frac{1}{x}\right)(D + x).$$

(d) Resuelva la ecuación diferencial

$$\left(D + \frac{1}{x}\right)(D + x)y = 0.$$

8. (a) Encuentre la solución general de la ecuación diferencial $y'' - y = 0$.

(b) Determine las constantes en la solución general $y(x)$ de la parte (a) de modo que $y(0) = 0$ e $y'(0) = 1$. Dibuje el gráfico de la solución particular obtenida así.

(c) Determine las constantes en la solución general $y(x)$ de la parte (a) de modo que se tenga que $y(0) = 1$ e $y(1) = 0$.
Dibuje el gráfico de la solución particular resultante.

9. (a) Demuestre que la ecuación característica de la ecuación diferencial $y'' + y = 0$ tiene raíces complejas $r_1 = i$ y $r_2 = -i$.

(b) Usando la definición

$$e^{\pm ix} = \cos x \pm i \operatorname{sen} x,$$

demuestre que la solución formal

$$y(x) = c_1e^{ix} + c_2e^{-ix}$$

de la ecuación diferencial $y'' + y = 0$ puede ser puesta en la forma

$$y(x) = d_1 \cos x + d_2 \operatorname{sen} x.$$

(c) Verifique directamente que cos x y sen x son soluciones de $y'' + y = 0$.

10. (Método de los coeficientes indeterminados.) Dada una ecuación diferencial *no homogénea*

$$(D^2 + aD + b)y = f(x), \tag{1}$$

supongamos que se sabe que la función $f(x)$ es una solución particular de una ecuación de una forma similar pero *homogénea*, digamos

$$(D - r_1) \ldots (D - r_n)f = 0.$$

Aplicando los operadores $(D - r_j)$ a ambos lados de la ecuación dada (1), se tiene que y también es una solución de la ecuación homogénea de orden más alto,

$$(D - r_1) \ldots (D - r_n)(D^2 + aD + b)y = 0. \tag{2}$$

Como tenemos un algoritmo para obtener la solución más general y_g de esta última ecuación, podemos encontrar una solución particular y_p de la ecuación dada (1) sustituyendo la solución general de (2) y considerando las condiciones resultantes para las constantes arbitrarias, o «coeficientes indeterminados», en y_g. Para evitar duplicaciones, podemos primero eliminar de y todos los términos que aparecen en la parte homogénea y_h de la solución de (1). La independencia lineal de las soluciones básicas del tipo de ecuación indicado garantiza que si se sustituye y_g en (1), entonces los coeficientes de términos análogos deben ser iguales. Entonces, la solución general de (1) es $y = y_h + y_p$. Encuentre las soluciones generales de las siguientes ecuaciones diferenciales.

(a) $y'' - y = e^{2x}$.

(b) $y'' - y = e^x$.

(c) $y'' + 2y' + y = e^x$.

(d) $y'' - y = x$.

(e) $y'' - y = e^x + x$.

11. (Variación de parámetros.) Este método es útil para encontrar soluciones particulares de ecuaciones diferenciales lineales con coeficientes no constantes de la forma

$$y'' + ay' + by = f(x). \tag{3}$$

Supongamos que podemos encontrar la solución homogénea

$$y_h = c_1u_1(x) + c_2u_2(x).$$

Las constantes c_1 y c_2 pueden ser consideradas como variables auxiliares o «parámetros» que pueden «variar» como función de x. Tratamos de determinar $c_1(x)$ y $c_2(x)$ de modo que

$$y(x) = c_1(x)u_1(x) + c_2(x)u_2(x) \tag{4}$$

sea una solución de la ecuación no homogénea (3). Usando la regla producto obtenemos

$$y' = c_1u_1' + c_2u_2' + c_1'u_1 + c_2'u_2$$

y luego, por simplicidad requerimos que

$$c_1'u_1 + c_2'u_2 = 0. \tag{5}$$

Por tanto,

$$y' = c_1u' + c_2u_2'. \tag{6}$$

A continuación calculamos

$$y'' = c_1u_1'' + c_2u_2'' + c_1'u_1' + c_2'u_2'. \tag{7}$$

Sustituyendo en la Ec. (3) las expresiones para y, y' e y'' suponiendo que u_1 y u_2 satisfacen la ecuación homogénea

$$y'' + ay' + by = 0,$$

se tiene que (véase Ejercicio 12)

$$c_1'u_1' + c_2'u_2' = f. \tag{8}$$

Resolviendo las Ecs. (5) y (8) tenemos que

$$c_1' = \frac{-u_2f}{u_1u_2' - u_2u_1'}, \qquad c_2' = \frac{u_1f}{u_1u_2' - u_2u_1'}. \tag{9}$$

Entonces podemos encontrar $c_1(x)$ y $c_2(x)$ por integración por lo cual (4) se convierte en una solución particular y_p que se debe sumar a y_h para obtener la solución general de (3).

Encuentre las soluciones generales de las siguientes ecuaciones no homogéneas.

(a) $y'' + y' - 2y = \log x$. (b) $y'' - y = e^{2x}$.

12. Suponga que u_1 y u_2 son soluciones de

$$y'' + ay' + by = 0.$$

(a) Verifique que si se sustituyen las expresiones (4), (6) y (7) del Ejercicio 11 en la ecuación

$$y'' + ay' + by = f$$

se obtiene la Ec. (8).

(b) Verifique que las Ecs. (5) y (8) del Ejercicio 11 tienen la solución dada por la Ec. (9).

13. Una bala disparada horizontalmente en un medio viscoso tiene un desplazamiento de su posición inicial dado por $x = x(t)$, donde t es el tiempo. Designando por puntos las derivaciones con respecto al tiempo, la velocidad de la bala en el instante t es $\dot{x}(t)$ y su aceleración es $\ddot{x}(t)$.

Partiendo de principios físicos se puede demostrar teóricamente que x satisface la ecuación diferencial

$$m\ddot{x} + k\dot{x} = 0,$$

donde m es la masa de la bala y k una constante positiva que depende de la viscosidad del medio.

(a) Demuestre que si el desplazamiento inicial es $x(0) = 0$ y la velocidad inicial es $\dot{x}(0) = v_0$, entonces para $t \geq 0$,

$$x(t) = \frac{v_0 m}{k}(1 - e^{-kt/m}).$$

(b) Demuestre que el desplazamiento $x(t)$ tiene una cota superior igual a $v_0 m/k$. ¿Cuál es el efecto de aumentar la constante de viscosidad o aumentar la masa de la bala?

(c) Demuestre que la velocidad de la bala tiende a cero a medida que t aumenta.

(d) Demuestre que la aceleración de la bala es negativa para $t \geq 0$. ¿Cuál es el efecto en la aceleración de un aumento de k o de un aumento de m?

(e) Dibuje el gráfico de $x(t)$ si $v_0 m/k = 1$.

14. Sea $D = d/dt$. El sistema de ecuaciones diferenciales de primer orden

$$(aD + \alpha)y + (bD + \beta)z = f(t)$$

$$(cD + \gamma)y + (dD + \delta)z = g(t)$$

contiene el sistema puramente algebraico

$$\alpha y + \beta z = f(t)$$

$$\gamma y + \delta z = g(t)$$

como un caso especial cuando a, b, c, d son todos cero. El método de eliminación puede ser usado para resolver tanto el sistema diferencial como el algebraico. Para encontrar $y(t)$ y $z(t)$ se opera con $(dD + \delta)$ en la primera ecuación y con $(bD + \beta)$ en la segunda. Luego se resta una ecuación de la otra. De la ecuación resultante se puede encontrar $y(t)$ ya que ella no contiene a z. Luego se sustituye la solución general $y(t)$ en *una* de las ecuaciones y se encuentra $z(t)$. Para determinar posibles relaciones entre las constantes de integración *puede* ser necesario sustituir $y(t)$ y $z(t)$ en la *otra* ecuación dada. A veces se pueden hacer simplificaciones en el procedimiento indicado.

(a) Use el método recién descrito para encontrar la solución general del sistema

$$(D + 1)y + z = 0,$$

$$3y + (D - 1)z = 0.$$

(b) Determine las constantes en la solución de la parte (a) necesarias para que se cumplan las condiciones iniciales $y(0) = 0$ y $z(0) = 1$.

(c) Encuentre la solución más general del sistema

$$(D + 1)y + Dz = 0$$

$$Dy - (D - 1)z = t.$$

(d) Determine las constantes en la solución de la parte (c) necesarias para que se cumplan las condiciones iniciales $y(0) = 0$ y $z(0) = 0$.

15. Suponga que dos estanques de 100 galones contienen soluciones de sal de concentraciones (en libras por galón) $y(t)$ y $z(t)$ respectivamente. Suponga que la solución fluye del estanque y al z con una velocidad de 1 galón por minuto y del estanque z al y con una velocidad de 4 galones por minuto y que el exceso del estanque y se pierde mientras que el estanque z se mantiene lleno agregando agua pura. Suponemos que las soluciones en ambos estanques se mantienen perfectamente homogéneas en todo instante.

(a) Demuestre que y y z satisfacen un sistema de ecuaciones diferenciales del tipo considerado en el Ejercicio 14. [*Indicación*: Exprese Dy y Dz como combinaciones lineales de y y z.]

(b) Encuentre la solución general del sistema encontrado en la parte (a) y determine las constantes consistentes con concentraciones iniciales $y(0) = \frac{1}{10}$, $z(0) = \frac{1}{20}$.

(c) Dibuje los gráficos de las soluciones particulares $y(t)$ y $z(t)$ encontradas en la parte (b) e interprete físicamente los resultados obtenidos.

7. Espacios vectoriales complejos

Hasta ahora, hemos considerado que en los espacios vectoriales estudiados la multiplicación numérica se hace siempre con números reales. Sin embargo, si remplazamos los números reales por números complejos, la definición de espacio vectorial y de función lineal no varía. Es más, todos los teoremas que hemos demostrado para espacios vectoriales reales siguen siendo verdaderos con respecto a los números complejos. Para demostrar esto basta notar que las únicas propiedades de los números reales que se usan en un espacio vectorial abstracto son satisfechas también por los números complejos. No sucede lo mismo con los teoremas que usan los productos internos, esos teoremas serán discutidos al final de esta sección. Como motivación para la consideración de los espacios vectoriales complejos, explicaremos cómo esta extensión es la base de un estudio ulterior de los operadores diferenciales, que hemos empezado en la Sec. 6. Primero revisaremos las propiedades relevantes de los números complejos y de las funciones complejas.

El número complejo $z = x + iy$ con **parte real** Re $z = x$ y **parte imaginaria** Im $z = y$, puede ser identificado con el vector (x, y) de $\mathcal{R}^2$ cuando se desea tener una representación pictórica. Es más, la definición de **suma compleja** está definida de modo que corresponde exactamente a la suma de elementos de $\mathcal{R}^2$:

$$(x + iy) + (x' + iy') = (x + x') + i(y + y').$$

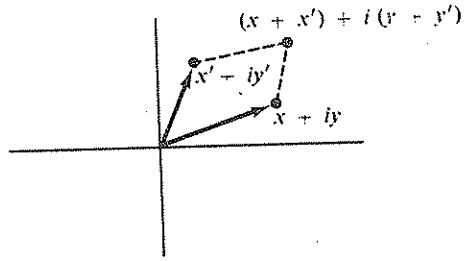


Figura 7

Esta relación se ilustra en la Fig. 7. También es verdad que la **multiplicación numérica** de un número complejo por un número **real** r corresponde a la multiplicación numérica en $\mathcal{R}^2$:

$$r(x + iy) = rx + iry.$$

Para completar la identificación de los números complejos con los vectores de $\mathcal{R}^2$ es necesario introducir una operación de multiplicación vectorial en $\mathcal{R}^2$ que corresponda a la multiplicación de números complejos. La **multiplicación compleja** se hace definiendo $i^2 = -1$ y calculando como sigue.

$$(x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx').$$

El **complejo conjugado** de $x + iy$, denotado por $\overline{x + iy}$ es

$$\overline{x + iy} = x - iy.$$

y la división por un número complejo $x + iy$ distinto de cero se efectúa simplemente multiplicando numerador y denominador por el conjugado del denominador:

$$\frac{x' + iy'}{x + iy} = \frac{(x' + iy')(x - iy)}{(x + iy)(x - iy)} = \frac{(xx' + yy') + i(xy' - yx')}{x^2 + y^2}.$$

Finalmente, el **valor absoluto** de $x + iy$, denotado por $|x + iy|$ se define como

$$|x + iy| = \sqrt{x^2 + y^2}$$

Nótese que el valor absoluto y el conjugado de un número complejo están relacionados por

$$|z|^2 = z\bar{z}.$$

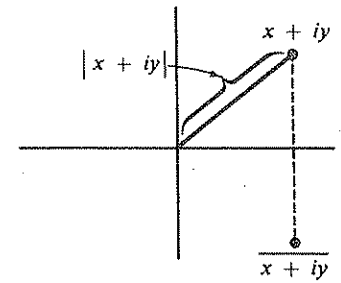


Figura 8

y estas dos definiciones corresponden, respectivamente, a encontrar el largo en $\mathcal{R}^2$ y a reflejar en el eje horizontal, como lo muestra la Fig. 8.

El significado geométrico de la multiplicación geométrica se comprende mejor si representamos los números complejos en su **forma polar:**

$$x + iy = |x + iy|\left(\frac{x}{|x + iy|} + i\frac{y}{|x + iy|}\right).$$

Como $|x + iy| = \sqrt{x^2 + y^2}$, los números $x/\sqrt{x^2 + y^2}$ y $y/\sqrt{x^2 + y^2}$ pueden ser expresados como el coseno y el seno, respectivamente, de un ángulo θ llamado el **ángulo polar de** $z = x + iy$. La Fig. 9 ilustra este ángulo. Es claro que todo número complejo distinto de cero tiene infinitos ángulos polares y cada par de ellos difieren en un múltiplo de 2π.

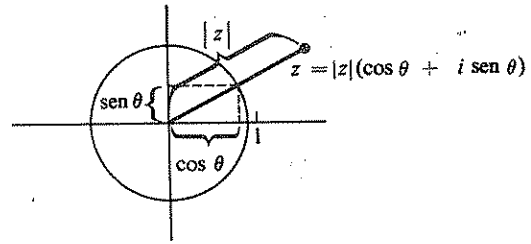


Figura 9

Ahora bien, si z y z' son números complejos con ángulos polares θ y θ', entonces podemos escribir su producto en forma polar como sigue:

$$\begin{aligned}&(|z|(\cos\theta + i\,\mathrm{sen}\,\theta))(|z|'(\cos\theta' + i\,\mathrm{sen}\,\theta'))\\ &\quad = |z|\,|z'|\,((\cos\theta\cos\theta' - \mathrm{sen}\,\theta\,\mathrm{sen}\,\theta') + i(\cos\theta\,\mathrm{sen}\,\theta' + \mathrm{sen}\,\theta\cos\theta'))\\ &\quad = |z|\,|z'|\,(\cos(\theta + \theta') + i\,\mathrm{sen}(\theta + \theta')).\end{aligned}$$

En el último paso hemos usado las fórmulas para el coseno y el seno de una suma. El resultado de este cálculo es un número en forma polar con valor abso-

luto $|z|\,|z'|$ y ángulo polar $\theta + \theta'$. Concluimos que el valor absoluto del producto de los números complejos z y z' es el producto de sus valores absolutos:

$$|zz'| = |z|\,|z'|,$$

y que si z y z' tienen ángulos polares $\theta(z)$ y $\theta(z')$ entonces zz' tiene ángulo polar $\theta(z) + \theta(z')$. La Fig. 10 ilustra estos hechos.

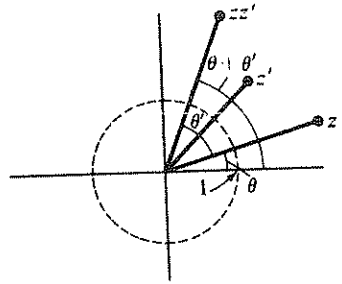


Figura 10

Lo que hemos hecho puede ser expresado en la forma más conveniente usando la función **exponencial compleja.** Definimos para un número real θ,

$$e^{i\theta} = \cos\theta + i\,\text{sen}\,\theta.$$

Por tanto, $e^{i\theta}$ es un número complejo con $|e^{i\theta}| = \sqrt{\text{sen}^2\,\theta + \cos^2\theta} = 1$ y con ángulo polar θ. Como los ángulos polares se suman al multiplicar números complejos, se tiene que

$$e^{i\theta}e^{i\theta'} = e^{i(\theta+\theta')}.$$

En particular, si $\theta' = -\theta$, tenemos que $e^{i\theta}e^{-i\theta} = 1$, de modo que

$$\frac{1}{e^{i\theta}} = e^{-i\theta} = \overline{e^{i\theta}}.$$

Estas últimas ecuaciones justifican el uso de la notación exponencial; esta función se comporta como la exponencial real para la cual $e^{\theta}e^{\theta'} = e^{\theta+\theta'}$ y $1/e^{\theta} = e^{-\theta}$.

En términos de la exponencial, la forma polar de un número complejo z con ángulo polar $\theta(z)$, toma la forma

$$z = |z|e^{i\theta(z)},$$

y su conjugado es

$$\bar{z} = |z|e^{-i\theta(z)}.$$

El hecho de que el conjugado de zw es el producto de los conjugados de z y w, puede ser demostrado directamente o como sigue:

$$\begin{aligned}\bar{z}\bar{w} &= |z|e^{-i\theta(z)}|w|e^{-i\theta(w)} \\ &= |zw|\,e^{-i(\theta(z)+\theta(w))} = \overline{zw}.\end{aligned}$$

Además de su simplicidad algebraica, otra razón para usar la notación exponencial compleja proviene de las fórmulas para su derivada e integral. Para derivar o integrar una función con valores complejos $u(x) + iv(x)$ con respecto de la variable real x, simplemente derivamos o integramos su parte real y su parte imaginaria. Por definición,

$$\frac{d}{dx}\big(u(x) + iv(x)\big) = \frac{du}{dx}(x) + i\frac{dv}{dx}(x),$$

y

$$\int\big(u(x) + iv(x)\big)\,dx = \int u(x)\,dx + i\int v(x)\,dx.$$

Entonces la derivada de e^{ix} con respecto de x está dada por

$$\begin{aligned}\frac{d}{dx}e^{ix} &= \frac{d}{dx}(\cos x + i\,\text{sen}\,x) \\ &= -\text{sen}\,x + i\cos x \\ &= i(\cos x + i\,\text{sen}\,x) = ie^{ix}.\end{aligned}$$

En resumen

$$\frac{d}{dx}e^{ix} = ie^{ix}.$$

En forma análoga,

$$\int e^{ix}\,dx = \frac{1}{i}e^{ix} + c,$$

donde c es una constante real o compleja. Estas fórmulas son análogas a las fórmulas para la derivada e integral de $e^{\alpha x}$ cuando a es real. Más generalmente podemos definir

$$e^{(\alpha+i\beta)x} = e^{\alpha x}e^{i\beta x}$$

y calcular

7.1
$$\frac{d}{dx}e^{(\alpha+i\beta)x} = (\alpha + i\beta)e^{(\alpha+i\beta)x}$$

y

7.2
$$\int e^{(\alpha+i\beta)x}\,dx = \frac{1}{\alpha + i\beta}e^{(\alpha+i\beta)x} + c, \qquad \alpha + i\beta \neq 0.$$

Estos cálculos se dejan como ejercicios.

Ahora podemos discutir la ecuación diferencial

$$(D^2 + aD + b)y = 0$$

cuando la factorización

$$(D - r_1)(D - r_2) = D^2 - (r_1 + r_2)D + r_1r_2$$

contiene números complejos r_1 y r_2. Veremos que se pueden aplicar las técnicas habituales discutidas, por ejemplo, en la Sec. 6. De hecho, el método del

multiplicador exponencial no cambia formalmente debido a la Ec. 7.1. Tenemos que

$$D(e^{rx}y) = e^{rx}(D + r)y,$$

donde r es un número real o complejo.

Ejemplo 1. Consideremos la ecuación diferencial $y'' + y = 0$. Usando operadores, esta ecuación tiene la forma

$$(D^2 + 1)y = 0,$$

factorizando $D^2 + 1$, tenemos

$$(D - i)(D + i)y = 0.$$

Hacemos

$$(D + i)y = u, \tag{1}$$

y tratamos de resolver en u la ecuación

$$(D - i)u = 0$$

Como en el caso real, multiplicamos ambos miembros por un factor tal que el miembro izquierdo sea la derivada de un producto. La misma regla del multiplicador sugiere que el factor adecuado es e^{-ix}, de modo que escribimos

$$e^{-ix}(D - i)u = 0$$

o, como $D(e^{-ix}u) = e^{-ix}(D - i)u$,

$$D(e^{-ix}u) = 0.$$

Integrando ambos miembros con respecto de x, obtenemos

$$e^{-ix}u = c_1 \quad \text{o} \quad u = c_1e^{ix}.$$

Sustituyendo este valor de u en la Ec. (1), obtenemos

$$(D + i)y = c_1e^{ix},$$

que ahora debe ser resuelto en y. Para ello multiplicamos por e^{ix} obteniendo

$$e^{ix}(D + i)y = c_1e^{2ix}$$

o, ya que $D(e^{ix}y) = e^{ix}(D + i)y$,

$$D(e^{ix}y) = c_1e^{2ix}.$$

Integrando tenemos que

$$e^{ix}y = \frac{1}{2i}c_1e^{2ix} + c_2$$

es decir,

$$y = \frac{1}{2i}c_1e^{ix} + c_2e^{-ix}.$$

Si remplazamos la constante arbitraria c_1 por $2ic_1$, tenemos

$$\begin{aligned} y &= c_1e^{ix} + c_2e^{-ix} \\ &= c_1(\cos x + i \operatorname{sen} x) + c_2(\cos x - i \operatorname{sen} x) \\ &= (c_1 + c_2)\cos x + i(c_1 - c_2)\operatorname{sen} x. \end{aligned}$$

Para darle una forma más sencilla a esta fórmula, hacemos $d_1 = (c_1 + c_2)$ y $d_2 = i(c_1 - c_2)$. Esto no implica una pérdida en la generalidad de las constantes, ya que para valores cualesquiera de d_1 y d_2, podemos despejar c_1 y c_2. Sin embargo, si d_1 y d_2 deben ser números reales, entonces c_1 y c_2 deben ser, en general, complejos. De hecho, tenemos que

$$c_1 = \frac{d_1 + id_2}{2} \quad \text{y} \quad c_2 = \frac{d_1 - id_2}{2}.$$

El ejemplo anterior muestra que es importante poder formar combinaciones lineales, con coeficientes complejos, en un espacio vectorial. Un **espacio vectorial complejo** tiene la misma definición que un espacio vectorial real excepto que los múltiplos numéricos se forman usando los números complejos $\mathbb{C}$. Las definiciones de independencia lineal y de función lineal son formalmente las mismas, con números reales remplazados por números complejos. Es conveniente notar que todo espacio vectorial complejo puede ser convertido en un espacio vectorial real de una manera natural: simplemente nos restringimos a la multiplicación numérica por números reales. Como consecuencia, la independencia lineal de un conjunto en un espacio vectorial complejo implica automáticamente la independencia lineal del mismo conjunto relativa al espacio vectorial real restringido. Esto se debe a que si

$$c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = 0$$

implica que todos los c son cero cuando se los toma como números complejos, entonces, con mayor razón, lo serán si se toman como números reales. Sin embargo, la afirmación conversa no es válida, como lo muestra el siguiente ejemplo.

Ejemplo 2. El conjunto $\mathbb{C}$ de los números complejos es un espacio vectorial complejo. Hemos visto que, como espacio vectorial real, $\mathbb{C}$ puede ser identificado con $\mathbb{R}^2$ y, por tanto, tiene dimensión 2 relativa a los números reales. Sin embargo, con respecto de los números complejos, $\mathbb{C}$ tiene dimensión 1. Esto se debe a que una combinación lineal con coeficientes complejos de más de un número complejo puede siempre ser hecha cero con coeficientes no todos cero. Por ejemplo, en $c_1z + c_2w$, elegimos $c_1 = 1/z$ y $c_2 = -1/w$.

Volviendo a las ecuaciones diferenciales, notamos que los Teoremas 6.1, 6.2 y 6.3 de la sección anterior son verdaderos con respecto a los números complejos. La razón para esto es que en las demostraciones de estos teoremas lo único que se presupone acerca de los multiplicadores numéricos son las leyes ordinarias de la aritmética que son válidas tanto para los números reales como para los números complejos. Además, las propiedades de la función exponencial, que se usan, también son válidas en los dos casos. Por tanto, concluimos que todas las soluciones de la ecuación diferencial $L(y) = 0$, donde

$$L = (D - r_1)\ldots(D - r_n) \tag{2}$$

son combinaciones lineales de n funciones, aun en el caso en que los números r_k son complejos. Estas funciones son e^{r_kx} o, en el caso de raíces múltiples, $x^ae^{r_kx}$. Esto implica, tal como para r_k real, que:

7.3 Teorema

El conjunto de soluciones de $L(y) = 0$, donde L está definido por la Ec. (2), es un espacio vectorial $\mathfrak{N}$ de dimensión n con respecto de los números complejos. $\mathfrak{N}$ tiene una base formada por funciones de la forma $e^{r_k x}$ o $x^q e^{r_k x}$, donde r_k puede ser complejo.

Generalmente, el Teorema 7.3 no se aplica directamente en la forma indicada anteriormente. La razón para esto es que los operadores que aparecen en la práctica, tales como

$$P(D) = D^2 + aD + b,$$

casi siempre tienen coeficientes reales a y b. Esto implica que en la factorización

$$D^2 + aD + b = (D - r_1)(D - r_2),$$

las raíces **complejas** siempre aparecen en pares conjugados. (Si r es una raíz de $P(x)$ con coeficientes **reales,** entonces $\bar{r}$ también es una raíz. ¿Por qué?) Por tanto, las soluciones de la ecuación diferencial también ocurren en pares de la forma

$$e^{rx}, \qquad e^{\bar{r}x},$$

posiblemente multiplicados por alguna potencia de x. Entonces es natural escribir $r = \alpha + i\beta$ y $\bar{r} = \alpha - i\beta$, con lo cual

$$\begin{aligned} c_1 e^{rx} + c_2 e^{\bar{r}x} &= c_1 e^{(\alpha+i\beta)x} + c_2 e^{(\alpha-i\beta)x} \\ &= e^{\alpha x}\big(c_1(\cos\beta x + i\,\mathrm{sen}\,\beta x) + c_2(\cos\beta x - i\,\mathrm{sen}\,\beta x)\big) \\ &= d_1 e^{\alpha x}\cos\beta x + d_2 e^{\alpha x}\,\mathrm{sen}\,\beta x. \end{aligned}$$

donde $d_1 = c_1 + c_2$ y $d_2 = i(c_1 - c_2)$. Las funciones $e^{\alpha x}\cos\beta x$ y $e^{\alpha x}\,\mathrm{sen}\,\beta x$ tienen interpretaciones geométricas más sencillas que la exponencial compleja. Por tanto, a menudo las soluciones se expresan usando la forma trigonométrica

$$y = d_1 e^{\alpha x}\cos\beta x + d_2 e^{\alpha x}\,\mathrm{sen}\,\beta x.$$

Ejemplo 3. La ecuación diferencial

$$y'' + 2y' + 2y = 0$$

tiene polinomio característico

$$D^2 + 2D + 2.$$

Las raíces de este polinomio son $-1 \pm i$, de modo que la ecuación puede ser escrita en la forma

$$(D - (-1 - i))(D - (-1 + i))y = 0.$$

Las soluciones exponenciales complejas son

$$e^{(-1-i)x}, \qquad e^{(-1+i)x}.$$

La forma geométrica de estas soluciones es

$$e^{-x}\cos x, \qquad e^{-x}\,\mathrm{sen}\,x;$$

de modo que la solución general puede ser escrita ya sea en la forma

$$c_1 e^{(-1-i)x} + c_2 e^{(-1+i)x}$$

o bien

$$d_1 e^{-x}\cos x + d_2 e^{-x}\,\mathrm{sen}\,x.$$

La contrapartida natural del Teorema 7.3 para ecuaciones diferenciales con coeficientes reales es el teorema siguiente que garantiza la existencia de una base para el espacio de las soluciones reales de $L(y) = 0$.

7.4 Teorema

El conjunto de las soluciones reales de

$$(D^n + a_{n-1}D^{n-1} + \ldots + a_1 D + a_0)y = 0,$$

donde los a_k son reales, es un espacio vectorial $\mathfrak{M}$ de dimensión n con respecto de los números reales. $\mathfrak{M}$ tiene una base formada por las funciones de la forma

$$x^q e^{\alpha x}\cos\beta x, \qquad x^q e^{\alpha x}\,\mathrm{sen}\,\beta x,$$

donde α y β son reales.

Demostración. Por el Teorema 7.3 sabemos que las soluciones complejas de la ecuación diferencial considerada son combinaciones lineales de funciones de la forma $x^q e^{rx}$, donde r es real o, en caso contrario, existe otra solución de la forma $x^q e^{\bar{r}x}$. Hemos visto que toda solución, en particular toda solución real, puede ser escrita como una función lineal de n funciones de la forma

$$x^q e^{\alpha x}\cos\beta x, \qquad x^q e^{\alpha x}\,\mathrm{sen}\,\beta x. \tag{3}$$

Como el espacio $\mathfrak{M}$ de las soluciones tiene dimensión n, cualquier conjunto de n funciones que lo generen deben ser linealmente independientes con respecto de los números complejos. Entonces, estas mismas n funciones serán automáticamente linealmente independientes con respecto de los números reales y, por tanto, forman una base de $\mathfrak{M}$.

Concluimos que n funciones de la forma (3) forman una base de $\mathfrak{M}$.

Terminamos esta sección con algunas consideraciones adicionales acerca de los espacios vectoriales complejos.

Ejemplo 4. Denotamos por $\mathbb{C}^n$ el conjunto de todas las n-plas de números complejos. Es fácil comprobar que $\mathbb{C}^n$ es un espacio vectorial con respecto de la suma y multiplicación numérica definida por coordenadas. Salvo indicación expresa, siempre subentenderemos que la multiplicación numérica en $\mathbb{C}^n$ es con respecto de los números complejos. Entonces $\mathbb{C}^n$ tiene la base $\mathbf{e}_1 = (1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, \ldots, 0, 1)$ y, por tanto, tiene dimensión n.

Ejemplo 5. Sea P_n el espacio vectorial de los polinomios en potencias de x de 1 a x^n, con coeficientes complejos. Entonces P_n tiene dimensión $n + 1$, puesto que el conjunto generador $\{1, x, \ldots, x^n\}$ es linealmente independiente. Para ver esto, notemos que

$$c_0 + c_1 x + \ldots + c_n x^n = 0 \quad \text{para todo } x$$

implica que el polinomio tiene más de n raíces. Esto solo es posible si todos los coeficientes son cero.

Es posible definir largo y producto interior (análogo del producto punto en $\mathcal{R}^n$) en un espacio vectorial complejo. Para ver cómo debe hacerse esto, supongamos que el producto interior de dos vectores $\mathbf{x}$ e $\mathbf{y}$ es la función con valores complejos $\langle \mathbf{x}, \mathbf{y} \rangle$. Supongamos también que $\langle \mathbf{x}, \mathbf{x} \rangle$ es no negativa, de modo que podamos definir el largo de $\mathbf{x}$ por $|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. Si suponemos, además, que $\langle \mathbf{x}, \mathbf{x} \rangle$ es lineal compleja en $\mathbf{x}$ y en $\mathbf{y}$, entonces para todo número complejo c, se tiene que

$$|c\mathbf{x}|^2 = \langle c\mathbf{x}, c\mathbf{x} \rangle = c^2 \langle \mathbf{x}, \mathbf{x} \rangle = c^2 |\mathbf{x}|^2.$$

Pero $|c\mathbf{x}|^2$ y $|\mathbf{x}|^2$ son números reales mientras que, en general, el cuadrado c^2 del número complejo c, no lo es. Para evitar esta dificultad, exigimos que $\langle \mathbf{x}, \mathbf{y} \rangle$ o sea **conmutatividad conjugada**

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle},$$

Por tanto, definimos un **producto interior complejo** $\langle \mathbf{x}, \mathbf{y} \rangle$ de los elementos $\mathbf{x}$, $\mathbf{y}$ de un espacio vectorial complejo de modo que las siguientes propiedades sean válidas.

7.5 Positividad: $\langle \mathbf{x}, \mathbf{x} \rangle > 0$, excepto por $\langle 0, 0 \rangle = 0$.

Conmutatividad conjugada: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$.

Distributividad: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.

Asociatividad: $\langle c\mathbf{x}, \mathbf{y} \rangle = c \langle \mathbf{x}, \mathbf{y} \rangle$.

La simetría conjugada implica la aditividad en el segundo vector, de modo que

$$\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle.$$

Sin embargo, tenemos la **asociatividad conjugada** en la segunda variable:

$$\langle \mathbf{x}, c\mathbf{y} \rangle = \overline{\langle c\mathbf{y}, \mathbf{x} \rangle} = \bar{c} \overline{\langle \mathbf{y}, \mathbf{x} \rangle} = \bar{c} \langle \mathbf{x}, \mathbf{y} \rangle.$$

Ejemplo 6. En el espacio vectorial complejo $\mathcal{C}^n$, sean $\mathbf{z} = (z_1, \ldots, z_n)$ y $\mathbf{w} = (w_1, \ldots, w_n)$. Definamos

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1 \overline{w}_1 + \ldots + z_n \overline{w}_n.$$

Es fácil verificar que esto define un producto interior complejo. Si definimos el largo de $\mathbf{z}$ por

$$|\mathbf{z}| = \langle \mathbf{z}, \mathbf{z} \rangle^{1/2},$$

entonces $|\mathbf{z}|$ tiene las propiedades del largo indicadas en 5.3 de la Sec. 5 del Cap. 1.

Ejemplo 7. Sea $\mathcal{E}_n$ el espacio vectorial de los polinomios exponenciales

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

definidos para $-\pi \leq x \leq \pi$ y con coeficientes complejos c_k. Podemos definir un producto interior complejo en $\mathcal{E}_n$ por

$$\langle p, q \rangle = \int_{-\pi}^{\pi} p(x) \overline{q(x)}\, dx.$$

La función compleja $p\bar{q}$ se integra simplemente integrando sus partes real e imaginaria. Definimos largo por

$$|p| = \langle p, p \rangle^{1/2} = \left(\int_{-\pi}^{\pi} |p(x)|^2\, dx \right)^{1/2}.$$

En cada uno de los ejemplos anteriores, el largo del vector complejo $\mathbf{z}$ se definió por $|\mathbf{z}| = \langle \mathbf{z}, \mathbf{z} \rangle^{1/2}$ usando un producto interior simétrico conjugado. Si consideramos las propiedades usuales del largo:

7.6 Positividad: $|\mathbf{z}| > 0$, excepto que $|0| = 0$.

Homogeneidad: $|c\mathbf{z}| = |c|\,|\mathbf{z}|$.

Desigualdad triangular: $|\mathbf{z} + \mathbf{w}| \leq |\mathbf{z}| + |\mathbf{w}|$,

Es obvio que la primera es válida en vista de la propiedad correspondiente del producto interno. La segunda se deduce de

$$|c\mathbf{z}| = \langle c\mathbf{z}, c\mathbf{z} \rangle^{1/2} = [\langle c\bar{c}\mathbf{z}, \mathbf{z} \rangle]^{1/2} = |c|\,|\mathbf{z}|.$$

Igual que en el caso de un producto interno real, y con una demostración igual, la desigualdad triangular se deduce de la desigualdad de Cauchy-Schwarz: $|\langle \mathbf{z}, \mathbf{w} \rangle| \leq |\mathbf{z}|\,|\mathbf{w}|$. La demostración de la desigualdad de Cauchy-Schwarz es una modificación sencilla de la demostración del caso real y aparece como el Ejercicio 11. La diferencia entre las demostraciones es una ilustración del hecho de que, como los productos internos complejos tienen propiedades algo distintas, no podemos esperar poder extender sin alteraciones los teoremas sobre productos interiores reales al caso complejo. Los productos interiores complejos se usan en este capítulo en los ejercicios de las Secs. 8 y 9.

EJERCICIOS

1. Para cada uno de los números complejos z que siguen, calcule $\bar{z}$ y $|z|$. Escriba z en forma polar y calcule $1/z$.
   (a) $1 + i$.
   (b) $-1 + 2i$.
   (c) $2i$.
   (d) $\dfrac{2 + i}{i}$.

2. Si z_1, z_2 y z_3 son números complejos, verifique que las leyes de distributividad

$$z_1(z_2 + z_3) = z_1z_2 + z_1z_3,$$

asociatividad

$$z_1(z_2z_3) = (z_1z_2)z_3 \quad \text{y} \quad z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3,$$

y conmutatividad

$$z_1z_2 = z_2z_1 \quad \text{y} \quad z_1 + z_2 = z_2 + z_1$$

son válidas.

3. (a) Compruebe la Ec. 7.1.
   (b) Compruebe la Ec. 7.2.

4. Demuestre directamente a partir de las definiciones de conjugado y valor absoluto que $\overline{z_1z_2} = \bar{z}_1\bar{z}_2$ y $|z_1z_2| = |z_1|\,|z_2|$.

5. Forme dos series infinitas a partir de la serie

$$\sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!}$$

separando los términos reales e imaginarios, use el resultado para justificar la definición de $e^{i\theta}$ por

$$e^{i\theta} = \cos\theta + i \operatorname{sen}\theta.$$

6. Demuestre que si r es un número complejo entonces $(D + r)$ es lineal como operador en el espacio vectorial complejo formado por las funciones continuamente diferenciables de la forma $u(x) + iv(x)$ donde x es una variable real.

7. Encuentre las soluciones generales de las siguientes ecuaciones diferenciales.
   (a) $(D^2 + 1)y = 1$.
   (b) $(D^2 + D + 1)y = 0$.
   (c) $y'' + y = \operatorname{sen} x$.
   (d) $y'' + y = 1$.
   (e) $y'' + y = \operatorname{sen} x + 1$.
   (f) $y'' + y = \tan x$.

8. Un resorte que vibra en un medio viscoso tiene, en el instante t, un desplazamiento $x(t)$ de su posición inicial. Si se denota por un punto la derivación con respecto de t, entonces, a partir de consideraciones físicas se puede obtener la ecuación diferencial,

$$m\ddot{x} + k\dot{x} + hx = 0,$$

donde m es la masa del resorte, k es una constante positiva que depende de la viscosidad del medio y h es una constante positiva que depende de la elasticidad del resorte.
   (a) Considerando las raíces del polinomio característico demuestre que $x(t)$ es oscilatorio o no dependiendo de si la desigualdad $4mh > k^2$ se satisface o no.

   (b) Suponga que m, h y k son iguales a 1 y encuentre la solución de la ecuación diferencial que satisface $x(0) = 0$ y $\dot{x}(0) = 3$.
   (c) Si en la parte (b) se toma $k = 2$, manteniendo iguales las otras condiciones, encuentre la solución correspondiente de la ecuación.
   (d) ¿Cuál es el desplazamiento máximo bajo las condiciones de la parte (b)? Demuestre que la oscilación tiende a cero a medida que aumenta t.
   (e) Dibuje el gráfico de la función desplazamiento bajo las condiciones de la parte (c). ¿Cuál es y en qué momento ocurre el desplazamiento máximo?

9. (a) Demuestre que ($a \cos cx + b \operatorname{sen} cx$), donde a, b y c son números reales, puede ser expresado en la forma $r \cos(cx - \theta)$, donde $r = \sqrt{a^2 + b^2}$ y θ es un ángulo tal que $\cos\theta = a/r$ y $\operatorname{sen}\theta = b/r$.
   (b) El resultado de la parte (a) es útil ya que muestra que una combinación lineal de $\cos cx$ y $\operatorname{sen} cx$ tiene como gráfico al gráfico de $\cos cx$ girado en un ángulo apropiado θ llamado **ángulo de fase** y multiplicado por el número r, llamado **amplitud**. Dibuje el gráfico de

$$\frac{1}{2}\cos 2x + \frac{1}{\sqrt{3}}\operatorname{sen} 2x$$

encontrando primero r y un valor apropiado de

10. Demuestre directamente que $e^{\alpha x}\cos\beta x$ y $e^{\alpha x}\operatorname{sen}\beta x$ son linealmente independientes con respecto de los reales, usando las fórmulas

$$\cos\beta x = \frac{e^{i\beta x} + e^{-i\beta x}}{2}, \qquad \operatorname{sen}\beta x = \frac{e^{i\beta x} - e^{-i\beta x}}{2},$$

y el hecho de que r^x y $\bar{r}^x$ son linealmente independientes respecto de los complejos cuando $r \neq \bar{r}$.

11. Demuestre la desigualdad de Cauchy-Schwarz

$$|\langle \mathbf{z}, \mathbf{w}\rangle| \leq |\mathbf{z}|\,|\mathbf{w}|$$

para productos complejos. [Igual que en el caso real, empiece suponiendo que $|\mathbf{z}| = |\mathbf{w}| = 1$. Exprese en forma polar

$$\langle \mathbf{z}, \mathbf{w}\rangle = |\langle \mathbf{z}, \mathbf{w}\rangle|\, e^{i\theta},$$

de modo que

$$\langle e^{-i\theta}\mathbf{z}, \mathbf{w}\rangle = |\langle \mathbf{z}, \mathbf{w}\rangle|.$$

Luego exprese $|e^{-i\theta}\mathbf{z} - \mathbf{w}|^2$ en términos del producto interior complejo.]

12. Demuestre que el conjunto $\mathbb{C}^n$ de las n-plas de números complejos tienen dimensión n con respecto de los números complejos y dimensión $2n$ con respecto de los números reales.

13. Demuestre que si un conjunto de funciones reales es linealmente independiente con respecto de los números reales entonces también es linealmente independiente con respecto de los números complejos. (La afirmación conversa es verdadera más generalmente y ha sido tratada en el texto.)

8. Bases ortonormales

Si $\mathcal{U}$ es un espacio vectorial con producto interno y $\mathcal{U}$ tiene una base

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n,$$

entonces a menudo es deseable que la base sea un conjunto **ortonormal.** Esto significa que los vectores distintos $\mathbf{u}_i$ y $\mathbf{u}_j$ son perpendiculares,

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0, \qquad i \neq j, \tag{1}$$

y que cada uno tiene largo 1:

$$|\mathbf{u}_i| = \sqrt{\mathbf{u}_i \cdot \mathbf{u}_i} = 1, \qquad i = 1, \ldots, n. \tag{2}$$

Estas condiciones son útiles, ya que si un vector $\mathbf{x}$ en $\mathcal{U}$ se expresa como una combinación lineal de los vectores básicos en la forma

$$\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_j\mathbf{u}_j + \ldots + c_n\mathbf{u}_n,$$

entonces los coeficientes c_j pueden ser calculados fácilmente. En efecto,

$$\begin{aligned}\mathbf{x} \cdot \mathbf{u}_j &= c_1\mathbf{u}_1 \cdot \mathbf{u}_j + \ldots + c_j\mathbf{u}_j \cdot \mathbf{u}_j + \ldots + c_n\mathbf{u}_n \cdot \mathbf{u}_j \\ &= 0 + \ldots + c_j + \ldots + 0,\end{aligned}$$

de modo que

$$c_j = \mathbf{x} \cdot \mathbf{u}_j. \tag{3}$$

Ejemplo 1. Los vectores (1, 0) y (0, 1) en $\mathcal{R}^2$ forman una base ortonormal. Lo mismo es válido para el par de vectores $(1/\sqrt{2}, 1/\sqrt{2})$ y $(-1/\sqrt{2}, 1/-2)$. En términos de esta última base podemos escribir

$$(x, y) = c_1\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) + c_2\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Para determinar c_1 y c_2 usamos la Ec. (3), obteniendo

$$c_1 = (x, y) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{(x + y)}{\sqrt{2}},$$

$$c_2 = (x, y) \cdot \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{(-x + y)}{\sqrt{2}}.$$

En particular, si $(x, y) = (1, 2)$, tenemos que $c_1 = 3/\sqrt{2}$ y $c_2 = 1/\sqrt{2}$, de modo que

$$(1, 2) = \frac{3}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Ejemplo 2. El conjunto de funciones

$$\cos x, \operatorname{sen} x, \ldots, \cos nx, \operatorname{sen} nx$$

definidas para $-\pi \leq x \leq \pi$ forman una base para el espacio vectorial $\mathcal{T}_n$ de las sumas trigonométricas de la forma

$$T(x) = \sum_{k=1}^{n} (a_n \cos kx + b_k \operatorname{sen} kx).$$

Un producto interno natural en $\mathcal{T}_n$ es

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx,$$

y, con respecto de este producto interior, esta base es ortonormal. En efecto, se tiene que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos kx \cos lx\,dx = \begin{cases}0, & k \neq l\\ 1, & k = l\end{cases}$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{sen} kx \operatorname{sen} lx\,dx = \begin{cases}0, & k \neq l\\ 1, & k = l\end{cases}$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos kx \operatorname{sen} lx\,dx = 0.$$

Por tanto, los coeficientes a_k y b_k están dados por

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} T(x) \cos kx\,dx,$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} T(x) \operatorname{sen} kx\,dx.$$

Los números a_k y b_k se llaman los **coeficientes k-ésimos de Fourier** de T y sus aplicaciones serán consideradas en la Sec. 5 del Cap. 5.

Como las bases ortonormales son más simples, es natural hacerse la siguiente pregunta. Dada una base de un espacio vectorial con producto interno, ¿es posible encontrar una base ortonormal? La respuesta es afirmativa y describiremos un procedimiento efectivo. Primero hacemos la siguiente observación.

8.1 Teorema

Si $\mathbf{u}_1, \ldots, \mathbf{u}_n$ es un conjunto ortonormal en el espacio vectorial $\mathcal{U}$, entonces estos vectores son linealmente independientes. Los coeficientes de la combinación lineal

$$\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n$$

son $c_j = \mathbf{x} \cdot \mathbf{u}_j$.

Demostración. Supongamos que existen constantes $c_1, \ldots, c_n$ tales que

$$c_1\mathbf{u}_1 + \ldots + c_j\mathbf{u}_j + \ldots + c_n\mathbf{u}_n = 0.$$

Si para algún j hacemos el producto punto con $\mathbf{u}_j$ de ambos miembros de esta ecuación obtenemos

$$c_1\mathbf{u}_1 \cdot \mathbf{u}_j + \ldots + c_j\mathbf{u}_j \cdot \mathbf{u}_j + \ldots + c_n\mathbf{u}_n \cdot \mathbf{u}_j = 0.$$

Pero $\mathbf{u}_k \cdot \mathbf{u}_j = 0$ si $k \neq j$ y $\mathbf{u}_j \cdot \mathbf{u}_j = 1$, de modo que $c_j = 0$. Como j es arbitrario, los $\mathbf{u}$ son linealmente independientes.

Ahora supongamos que $\mathbf{x}_1, \ldots, \mathbf{x}_n$ es un conjunto linealmente independiente en el espacio vectorial $\mathcal{V}$; describimos un proceso para construir un conjunto ortonormal $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Primero elegimos uno cualquiera de los $\mathbf{x}$, digamos $\mathbf{x}_1$ y tomamos

$$\mathbf{u}_1 = \frac{\mathbf{x}_1}{|\mathbf{x}_1|}.$$

Entonces $|\mathbf{u}_1| = |\mathbf{x}_1|\,|\mathbf{x}_1| = 1$. Ahora elegimos otro de los $\mathbf{x}$, digamos $\mathbf{x}_2$ y consideramos su proyección en $\mathbf{u}_1$, es decir, consideramos el vector $(\mathbf{x}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$. Si los vectores fueran los vectores geométricos ordinarios, la relación entre $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{u}_1$ sería como la mostrada en la Fig. 11.

El vector $\mathbf{y}_2$ está definido por

$$\mathbf{y}_2 = \mathbf{x}_2 - (\mathbf{x}_2 \cdot \mathbf{u}_1)\mathbf{u}_1,$$

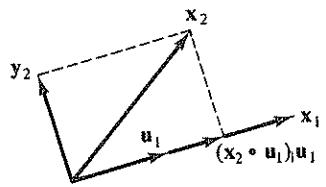


Figura 11

y podemos comprobar que $\mathbf{y}_2$ es perpendicular a $\mathbf{u}_1$. En efecto,

$$\begin{aligned}\mathbf{y}_2 \cdot \mathbf{u}_1 &= \mathbf{x}_2 \cdot \mathbf{u}_1 - (\mathbf{x}_2 \cdot \mathbf{u}_1)(\mathbf{u}_1 \cdot \mathbf{u}_1)\\ &= \mathbf{x}_2 \cdot \mathbf{u}_1 - \mathbf{x}_2 \cdot \mathbf{u}_1 = 0.\end{aligned}$$

Para obtener un vector *unitario* perpendicular a $\mathbf{u}_1$, hacemos

$$\mathbf{u}_2 = \frac{\mathbf{y}_2}{|\mathbf{y}_2|}.$$

El vector $\mathbf{y}_2$ no puede ser cero, ya que, por su definición, esto implicaría que $\mathbf{x}_2$ y $\mathbf{u}_1$ son linealmente dependientes.

Una vez encontrado $\mathbf{u}_1$ y $\mathbf{u}_2$, elegimos $\mathbf{x}_3$ y formamos su proyección en el subespacio de $\mathcal{V}$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Por definición, este es el vector

$$\mathbf{p} = (\mathbf{x}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}_3 \cdot \mathbf{u}_2)\mathbf{u}_2.$$

Definimos $\mathbf{y}_3$ restando esta proyección a $\mathbf{x}_3$.

$$\mathbf{y}_3 = \mathbf{x}_3 - (\mathbf{x}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{x}_3 \cdot \mathbf{u}_2)\mathbf{u}_2.$$

Como antes, podemos verificar que $\mathbf{y}_3$ es perpendicular a $\mathbf{u}_1$ y a $\mathbf{u}_2$. En efecto, como $\mathbf{u}_1 \cdot \mathbf{u}_1 = 1$ y $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, tenemos que

$$\mathbf{y}_3 \cdot \mathbf{u}_1 = (\mathbf{x}_3 \cdot \mathbf{u}_1) - (\mathbf{x}_3 \cdot \mathbf{u}_1) = 0.$$

Análogamente,

$$\mathbf{y}_3 \cdot \mathbf{u}_2 = (\mathbf{x}_3 \cdot \mathbf{u}_2) - (\mathbf{x}_3 \cdot \mathbf{u}_2) = 0.$$

En la Fig. 12 se muestran estos vectores como flechas. Definimos $\mathbf{u}_3$ por

$$\mathbf{u}_3 = \frac{\mathbf{y}_3}{|\mathbf{y}_3|}.$$

Nuevamente, $\mathbf{y}_3 = 0$ implicaría la dependencia lineal de $\mathbf{x}_3$, $\mathbf{u}_1$ y $\mathbf{u}_2$. Pero como

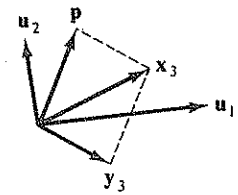


Figura 12

el subespacio generado por $\mathbf{u}_1$ y $\mathbf{u}_2$ es igual al generado por $\mathbf{x}_1$ y $\mathbf{x}_2$, esto implicaría la dependencia lineal de los $\mathbf{x}$.

Procediendo de esta manera, calculamos sucesivamente $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_j$. Para encontrar $\mathbf{u}_{j+1}$, escribimos

$$\mathbf{y}_{j+1} = \mathbf{x}_{j+1} - (\mathbf{x}_{j+1} \cdot \mathbf{u}_1)\mathbf{u}_1 - \ldots - (\mathbf{x}_{j+1} \cdot \mathbf{u}_j)\mathbf{u}_j,$$

8.2

$$\mathbf{u}_{j+1} = \frac{\mathbf{y}_{j+1}}{|\mathbf{y}_{j+1}|}.$$

Como anteriormente, podemos verificar que $\mathbf{u}_{j+1}$ es perpendicular a $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{j+1}$. Las Ecs. 8.2 resumen el así llamado **procedimiento de Gram-Schmidt** para encontrar un conjunto ortonormal a partir de un conjunto independiente. Se puede continuar hasta que se agoten los x.

El vector

$$(\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + \ldots + (\mathbf{x} \cdot \mathbf{u}_j)\mathbf{u}_j$$

se llama la **proyección** de $\mathbf{x}$ en el subespacio generado por $\mathbf{u}_1, \ldots, \mathbf{u}_j$. (La proyección se llama también el **desarrollo de Fourier** de $\mathbf{x}$ con respecto de $\mathbf{u}_1, \ldots, \mathbf{u}_j$. En el Teorema 7.1 del Cap. 5 se demuestra que el desarrollo de Fourier de $\mathbf{x}$ es la combinación lineal de los $\mathbf{u}$ más cercana a $\mathbf{x}$.)

Ejemplo 3. Los vectores $\mathbf{x}_1 = (1, -1, 2)$ y $\mathbf{x}_2 = (1, 0, -1)$ generan un plano P en $\mathcal{R}^3$, ya que ellos son linealmente independientes. Para encontrar

una base ortonormal de P aplicamos el procedimiento de Gram-Schmidt a $\mathbf{x}_1$ y $\mathbf{x}_2$. Tomamos

$$\mathbf{u}_1 = \frac{\mathbf{x}_1}{|\mathbf{x}_1|} = \frac{(1, -1, 2)}{\sqrt{6}} = \left(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right).$$

y calculamos

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{x}_2 - (\mathbf{x}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 \\ &= (1, 0, -1) - \left(\frac{-1}{\sqrt{6}}\right)\left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) \\ &= (1, 0, -1) + (\tfrac{1}{6}, -\tfrac{1}{6}, \tfrac{1}{3}) = (\tfrac{7}{6}, -\tfrac{1}{6}, -\tfrac{2}{3}).\end{aligned}$$

Luego

$$\begin{aligned}\mathbf{u}_2 &= \frac{\mathbf{y}_2}{|\mathbf{y}_2|} = \frac{(\frac{7}{6}, -\frac{1}{6}, -\frac{2}{3})}{\sqrt{66/36}} \\ &= \frac{1}{\sqrt{66}}(7, -1, -4).\end{aligned}$$

Por tanto, el plano P puede ser representado como el conjunto de todas las combinaciones lineales

$$u\mathbf{x}_1 + v\mathbf{x}_2$$

o de todas las combinaciones lineales

$$s\mathbf{u}_1 + t\mathbf{u}_2.$$

La relación entre los dos pares de vectores se muestra en la Fig. 13.

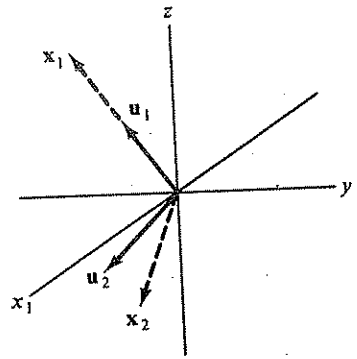


Figura 13

Ejemplo 4. Sea $P_n[-1, 1]$ el espacio vectorial de todos los polinomios $f(x) = a_0 + a_1x + \ldots + a_nx^n$ definidos para $-1 \leq x \leq 1$. Definimos un producto interno por

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx.$$

En el Ejemplo 5(b) de la Sec. 4, vimos que las funciones $1, x, \ldots, x^n$ forman una base de P_n. Para encontrar una base ortonormal, notamos que $\langle 1, 1\rangle = 2$, de modo que tomamos $u_1(x) = 1/\sqrt{2}$. Entonces, puesto que

$$y_2(x) = x - \left\langle x, \frac{1}{\sqrt{2}}\right\rangle \frac{1}{\sqrt{2}} = x,$$

puesto que $\langle x, 1/\sqrt{2}\rangle = \int_{-1}^{1} (x/\sqrt{2})\,dx = 0$. Entonces calculamos

$$\begin{aligned}u_2(x) &= \frac{x}{\langle x, x\rangle^{1/2}} = \frac{x}{\sqrt{\int_{-1}^{1} x^2\,dx}} \\ &= \sqrt{\tfrac{3}{2}}x.\end{aligned}$$

Ahora, tomamos

$$\begin{aligned}y_3(x) &= x^2 - \langle x^2, u_1(x)\rangle u_1(x) - \langle x^2, u_2(x)\rangle u_2(x) \\ &= x^2 - \left(\int_{-1}^{1} \frac{x^2}{\sqrt{2}}\,dx\right)\frac{1}{\sqrt{2}} - 0 \\ &= x^2 - \tfrac{1}{3}.\end{aligned}$$

Entonces

$$\begin{aligned}u_3(x) &= \frac{x^2 - \frac{1}{3}}{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3}\rangle^{1/2}} = \frac{x^2 - \frac{1}{3}}{\sqrt{\int_{-1}^{1} (x^2 - \frac{1}{3})^2\,dx}} \\ &= \sqrt{\tfrac{4.5}{8}}(x^2 - \tfrac{1}{3}) = \sqrt{\tfrac{5}{8}}(3x^2 - 1).\end{aligned}$$

Este proceso puede ser continuado indefinidamente. (Los polinomios resultantes se llaman polinomios de Legendre normalizados. En la Sec. 7 del Cap. 5 daremos otro método para calcularlos.)

Evidentemente, si partimos de dos bases para un espacio vectorial $\mathcal{U}$ y les aplicamos el procedimiento de Gram-Schmidt, obtendremos, en general, dos bases ortonormales distintas. En particular, si una de las dos bases dadas es ortonormal y aplicamos el procedimiento de Gram-Schmidt a la otra, podremos obtener una base distinta. El teorema que sigue da una condición bajo la cual las dos bases ortonormales resultantes son iguales, salvo orientación. Es necesario suponer que $\mathcal{U}$ es un espacio vectorial con respecto de la multiplicación escalar por números reales.

8.3 Teorema

Sean $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ dos bases ortonormales del espacio vectorial $\mathcal{V}$. Si los subespacios generados por $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ son iguales para $k = 1, 2, \ldots, n$, entonces $\mathbf{u}_k = \pm\mathbf{v}_k$ para $k = 1, 2, \ldots, n$.

Demostración. Sea $\mathbf{v}_k$ uno de los $\mathbf{v}$. Como, por hipótesis, $\mathbf{v}_k$ está en el subespacio generado por $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, tenemos que

$$\mathbf{v}_k = \sum_{j=1}^{k} r_j \mathbf{u}_j$$

para ciertos números reales r_j. Pero $\mathbf{v}_k$ es ortogonal a cada uno de $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$, ya que, por hipótesis, él es ortogonal a $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ y estos dos conjuntos generan el mismo subespacio. Si formamos el producto punto de ambos miembros de la ecuación anterior por $\mathbf{u}_j$, para $1 \leq j \leq k$, tenemos que

$$0 = \mathbf{v}_k \cdot \mathbf{u}_j = r_j, \qquad j = 1, \ldots, k-1,$$
$$\mathbf{v}_k \cdot \mathbf{u}_k = r_k.$$

Entonces $\mathbf{v}_k = (\mathbf{v}_k \cdot \mathbf{u}_k)\mathbf{u}_k$. Como los $\mathbf{u}$ y los $\mathbf{v}$ tienen largo 1, debemos tener que

$$|\mathbf{v}_k \cdot \mathbf{u}_k| = 1.$$

De esto se deduce que $\mathbf{v}_k \cdot \mathbf{u}_k = \pm 1$, de modo que $\mathbf{v}_k = \pm\mathbf{u}_k$.

Ejemplo 5. Consideremos la base natural $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$ y la base $\mathbf{x}_1 = (3, 0, 0)$, $\mathbf{x}_2 = (1, -1, 0)$, $\mathbf{x}_3 = (1, 1, 1)$ de $\mathcal{R}^3$. Es claro que $\mathbf{e}_1$ y $\mathbf{x}_1$ generan el mismo subespacio de $\mathcal{R}^3$. Los pares $\{\mathbf{e}_1, \mathbf{e}_2\}$ y $\{\mathbf{x}_1, \mathbf{x}_2\}$ generan el plano xy y ambas bases generan $\mathcal{R}^3$. Por el teorema anterior, si aplicamos el procedimiento de Gram-Schmidt a $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, obtenemos una base ortonormal de la forma $\{\pm\mathbf{e}_1, \pm\mathbf{e}_2, \pm\mathbf{e}_3\}$. Dejamos como ejercicio comprobar que obtenemos $\{\mathbf{e}_1, -\mathbf{e}_2, \mathbf{e}_3\}$. La Fig. 14 muestra los vectores.

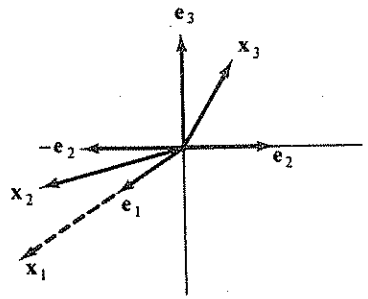


Figura 14

EJERCICIOS

1. (a) Encuentre un vector (x, y, z) tal que los vectores $(1, 1, 1)$, $(-1, \frac{1}{2}, \frac{1}{2})$, (x, y, z) formen una base ortogonal de $\mathcal{R}^3$.
   (b) Normalice la base obtenida en la parte (a).
2. Los vectores $(1, 1, 1)$ y $(1, 2, 1)$ generan un plano P en $\mathcal{R}^3$. Encuentre una base ortonormal para P usando el procedimiento de Gram-Schmidt.
3. Los vectores $(1, 2, 1, 1)$, $(-1, 0, 1, 0)$ y $(0, 1, 0, 2)$ forman una base de un subespacio $\mathcal{S}$ de $\mathcal{R}^4$. Use el procedimiento de Gram-Schmidt para encontrar una base ortonormal de $\mathcal{S}$.
4. Sea P_2 el espacio tridimensional de los polinomios cuadráticos $f(x)$ definidos para $0 \leq x \leq 1$. Si se considera el producto interno

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx,$$

de P_2, encuentre una base ortonormal de P_2. [*Indicación*: $\{1, x, x^2\}$ es una base de P_2.]

5. Demuestre que aplicando el procedimiento de Gram-Schmidt a los vectores $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ del Ejemplo 5 se obtienen $\mathbf{e}_1$, $-\mathbf{e}_2$, $\mathbf{e}_3$.
6. Demuestre que si se aplica el procedimiento de Gram-Schmidt a un conjunto ortonormal se obtiene el mismo conjunto.
7. Sea $C[-\pi, \pi]$ el espacio vectorial de las funciones complejas $f(x)$ definidas para $-\pi \leq x \leq \pi$. Defina

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx.$$

   para f y g en $C[-\pi, \pi]$.
   (a) Demuestre que $\langle f, g \rangle$ es un producto interior complejo.
   (b) Demuestre que

$$\langle e^{imx}, e^{inx} \rangle = \begin{cases} 1, & m = n. \\ 0, & m \neq n. \end{cases}$$

   (c) Demuestre que los subespacios generados por los siguientes conjuntos son iguales

$$\cos kx, \quad \operatorname{sen} kx, \quad k = 0, 1, 2, \ldots, n. \tag{1}$$
$$e^{ikx}, \quad k = 0, \quad \pm 1, \pm 2, \ldots, \pm n. \tag{2}$$

8. La conclusión del Teorema 8.3 es que $\mathbf{u}_k \pm \mathbf{v}_k$. ¿Qué conclusión es posible si $\mathcal{V}$ es un espacio vectorial complejo? [Véase el Ejercicio 11 de la Sec. 7.]

9. Vectores propios

En esta sección estudiaremos una manera natural de asociar una base de un espacio vectorial $\mathcal{V}$ a una función lineal f de $\mathcal{V}$ en $\mathcal{V}$. Supongamos que existe un vector $\mathbf{x}$ *distinto de cero* en $\mathcal{V}$ y un número λ tal que

$$f(\mathbf{x}) = \lambda\mathbf{x}.$$

Entonces $\mathbf{x}$ se llama un **vector propio** de la función lineal f y λ se llama su **valor propio** asociado. (Los términos **vector característico** y **raíz característica** se usan a veces.) Como f es lineal, la ecuación anterior siempre tendrá la solución trivial $\mathbf{x} = 0$ para todo λ. Esta solución no es interesante y no la consideraremos. También por linealidad, es claro que si $\mathbf{x}$ es un vector propio que corresponde a λ también lo es $c\mathbf{x}$ para todo número $c \neq 0$.

Ejemplo 1. Sea f la función lineal de $\mathcal{R}^2$ en $\mathcal{R}^2$ definida por la matriz

$$\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}.$$

Entonces

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ 4x+y \end{pmatrix}.$$

Es fácil verificar que

$$\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

y que

$$\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Por tanto, el vector $(1, 2)$ en $\mathcal{R}^2$ es un vector propio que corresponde al valor propio 3 y el vector $(1, -2)$ es un vector propio que corresponde al valor propio -1. Es claro que todo múltiplo numérico de estos vectores será un vector propio con el mismo valor propio que el vector original.

Antes de discutir cómo encontrar vectores propios, veremos por qué son útiles. Supongamos que $\mathcal{U}$ es un espacio vectorial y que f es una función lineal de $\mathcal{U}$ en $\mathcal{U}$. Supongamos, además, que $\mathcal{U}$ tiene una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ formada por vectores propios; es decir,

$$f(\mathbf{x}_k) = \lambda_k \mathbf{x}_k, \qquad k = 1, 2, \ldots, n,$$

para ciertos números λ_k. Entonces, si representamos un elemento arbitrario de $\mathcal{U}$ mediante esta base, f toma una forma muy simple. En efecto, si

$$\mathbf{x} = c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n.$$

Entonces, usando la linealidad de f y el hecho de que los $\mathbf{x}$ son vectores propios, tenemos que

$$\begin{aligned} f(\mathbf{x}) &= c_1 f(\mathbf{x}_1) + \ldots + c_n f(\mathbf{x}_n) \\ &= c_1\lambda_1\mathbf{x}_1 + \ldots + c_n\lambda_n\mathbf{x}_n. \end{aligned}$$

En otras palabras, con respecto de una base de vectores propios, la acción de

f consiste en multiplicar cada coeficiente por el valor propio correspondiente. Resumiendo,

$$f\left(\sum_{k=1}^{n} c_k\mathbf{x}_k\right) = \sum_{k=1}^{n} \lambda_k c_k \mathbf{x}_k, \tag{9.1}$$

donde

$$f(\mathbf{x}_k) = \lambda_k\mathbf{x}_k, \qquad k = 1, 2, \ldots, n. \tag{9.2}$$

En términos geométricos, si se usa una base de vectores propios, la acción de f puede ser interpretada como una sucesión de multiplicaciones numéricas, una para cada vector básico.

Ejemplo 2. Volviendo a la función lineal del Ejemplo 1, expresamos un vector arbitrario en $\mathcal{R}^2$ como una combinación lineal de los vectores propios $\mathbf{x}_1 = (1, 2)$ y $\mathbf{x}_2 = (1, -2)$:

$$\mathbf{x} = u\begin{pmatrix} 1 \\ 2 \end{pmatrix} + v\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Los valores propios correspondientes son 3 y -1, de modo que, partiendo de

$$f(\mathbf{x}) = uf\begin{pmatrix} 1 \\ 2 \end{pmatrix} + vf\begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

obtenemos

$$f(\mathbf{x}) = 3u\begin{pmatrix} 1 \\ 2 \end{pmatrix} - v\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

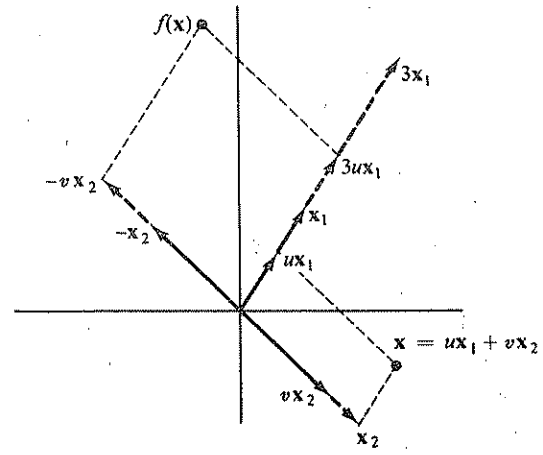


Figura 15

La Fig. 15 muestra el efecto de f en los dos vectores propios $\mathbf{x}_1$ y $\mathbf{x}_2$. Tenemos, entonces, que la imagen $f(\mathbf{x})$ de un vector cualquiera $\mathbf{x}$ puede ser construida por métodos geométricos. Además, podemos expresar el efecto f en palabras diciendo que f es la combinación de dos transformaciones:

1. Expandir a partir de la recta a lo largo de $\mathbf{x}_2$ en las rectas paralelas a $\mathbf{x}_1$.
2. Reflejar en la recta que contiene a $\mathbf{x}_1$ a lo largo de rectas paralelas a $\mathbf{x}_2$.

Para hacer este análisis es indispensable que los vectores propios de f generen a $\mathcal{R}^2$.

Ejemplo 3. Para encontrar los vectores propios y sus valores propios correspondientes de la función f de los ejemplos anteriores (si no los conociéramos), procedemos como sigue. Debemos encontrar vectores $\mathbf{x} \neq 0$ y números λ tales que

$$f(\mathbf{x}) - \lambda\mathbf{x} = 0.$$

En forma matricial, esta ecuación es

$$\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} - \lambda\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

o bien

$$\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

o, lo que es lo mismo,

$$\begin{pmatrix} (1-\lambda) & 1 \\ 4 & (1-\lambda) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{1}$$

Es claro que si la matriz 2-por-2 indicada tiene una inversa, entonces las únicas soluciones son $x = 0$, $y = 0$. Por tanto, debemos tratar de encontrar valores de λ tales que esa matriz *no* tenga una inversa. Esto ocurrirá precisamente si

$$\det\begin{pmatrix} (1-\lambda) & 1 \\ 4 & (1-\lambda) \end{pmatrix} = 0,$$

Es decir, si λ satisface $(1-\lambda)^2 - 4 = 0$. Esta ecuación cuadrática en λ tiene soluciones $\lambda = 3$ y $\lambda = -1$. Para encontrar los vectores propios correspondientes a $\lambda = 3$ y $\lambda = 1$, debemos encontrar x e y tales que ambos no son cero y que satisfagan la Ec. (1). Por tanto, consideramos

$$\lambda = 3: \quad \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y

$$\lambda = -1: \quad \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Estas ecuaciones se reducen a

$$\lambda = 3: \qquad -2x + y = 0$$
$$\lambda = -1: \qquad 2x + y = 0.$$

Por tanto, hay muchas soluciones, pero solo necesitamos una para cada valor propio. Por simplicidad, elegimos

$$\lambda = 3: \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\lambda = -1: \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

El método del ejemplo anterior puede ser resumido como sigue. Para encontrar los valores propios de una función lineal f de $\mathcal{R}^n$ en $\mathcal{R}^n$, resolvemos la **ecuación característica**

9.3 $$\det(A - \lambda I) = 0,$$

donde A es la matriz de f. Entonces, para cada valor propio $\lambda_1, \ldots, \lambda_n$, tratamos de encontrar vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ distintos de cero que satisfagan la ecuación

9.4 $$(A - \lambda_k I)\mathbf{x}_k = 0.$$

Como las Ecs. 9.3 y 9.4 están expresadas en términos de la matriz A de f, a menudo nos referimos a los vectores y valores propios de la *matriz* A. Es claro que la Ec. 9.4 es simplemente la Ec. 9.2 en forma matricial. Sin embargo, la Ec. 9.2 es más general, ya que se aplica también a funciones lineales no representables por matrices.

Ejemplo 4. Este ejemplo muestra que los vectores propios de una función lineal f no forman necesariamente una base para el espacio vectorial en el cual actúa f. Consideremos la función lineal f de $\mathcal{R}^2$ en $\mathcal{R}^2$ definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

donde θ es un número fijo, $\theta \leq \theta \leq 2\pi$. Esta función lleva

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \end{pmatrix}$$

y

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{pmatrix}$$

de modo que representa una rotación en un ángulo θ alrededor del origen en el sentido contrario al de los punteros del reloj. La Ec. 9.3 toma la forma

$$\det \begin{pmatrix} \cos\theta - \lambda & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta - \lambda \end{pmatrix} = 0$$

es decir,

$$\lambda^2 - 2\lambda\cos\theta + 1 = 0.$$

Esta ecuación en λ tiene raíces complejas conjugadas de la forma

$$\cos\theta \pm i\operatorname{sen}\theta = e^{\pm i\theta}.$$

Como estamos buscando vectores $\mathbf{x} \neq 0$ en $\mathcal{R}^2$ que satisfagan

$$f(\mathbf{x}) = e^{\pm i\theta}\mathbf{x},$$

concluimos que f solo tiene vectores propios en $\mathcal{R}^2$ si $\theta = 0$ o $\theta = \pi$. En el primer caso, tenemos la transformación identidad para la cual todos los vectores distintos de cero son vectores propios con valores propios 1. En el segundo caso, tenemos que $f(\mathbf{x}) = -\mathbf{x}$, que es una reflexión en el origen, y todos los vectores son vectores propios con valor propio -1. Es fácil darse cuenta por qué, en general, las rotaciones de $\mathcal{R}^2$ no tienen vectores propios. En el Ejercicio 8 se da un ejemplo de una función lineal que no define una base de vectores propios por razones distintas.

Terminaremos considerando la posibilidad de elegir vectores propios de una función lineal f de $\mathcal{U}$ en $\mathcal{U}$, de modo que ellos formen una base ortonormal. Como los vectores propios están determinados, salvo un múltiplo numérico, siempre los podemos normalizar. Podemos lograr la ortogonalidad si la función f es **simétrica** relativa a un producto interno en $\mathcal{U}$, es decir, si

$$f(\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot f(\mathbf{y})$$

para todo $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{U}$. En particular, si f es una función lineal de $\mathcal{R}^n$ en $\mathcal{R}^n$ con matriz $A = (a_{ij})$, entonces la condición de simetría es

$$A\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot A\mathbf{y}.$$

Ejemplo 5. Si f es una función lineal de $\mathcal{R}^2$ en $\mathcal{R}^2$ con matriz

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

simétrica con respecto de la diagonal, entonces f es una transformación simétrica. En efecto, tenemos que

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} y_1 \\ x_2 \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
$$= ax_1y_1 + b(x_2y_1 + x_1y_2) + cx_2y_2.$$

Análogamente,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\cdot\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\cdot\begin{pmatrix} ay_1 + by_2 \\ by_1 + cy_2 \end{pmatrix}$$
$$= ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2.$$

Como los dos productos punto son iguales, concluimos que f es simétrica.

9.5 Teorema

Sea f una función lineal simétrica de $\mathcal{U}$ en $\mathcal{U}$, donde $\mathcal{U}$ es un espacio vectorial con un producto interno. Si $\mathbf{x}_1$ y $\mathbf{x}_2$ son vectores propios de f con valores propios λ_1 y λ_2 distintos, entonces $\mathbf{x}_1$ y $\mathbf{x}_2$ son ortogonales.

Demostración. Debemos demostrar que $\mathbf{x}_1\cdot\mathbf{x}_2 = 0$ si $f(\mathbf{x}_1) = \lambda_1\mathbf{x}_1$ y $f(\mathbf{x}_2) = \lambda_2\mathbf{x}_2$. Tenemos que

$$f(\mathbf{x}_1)\cdot\mathbf{x}_2 = (\lambda_1\mathbf{x}_1)\cdot\mathbf{x}_2 = \lambda_1(\mathbf{x}_1\cdot\mathbf{x}_2)$$

y

$$\mathbf{x}_1\cdot f(\mathbf{x}_2) = \mathbf{x}_1\cdot(\lambda_2\mathbf{x}_2) = \lambda_2(\mathbf{x}_1\cdot\mathbf{x}_2).$$

Como f es simétrica, $f(\mathbf{x}_1)\cdot\mathbf{x}_2 = \mathbf{x}_1\cdot f(\mathbf{x}_2)$, de modo que

$$\lambda_1(\mathbf{x}_1\cdot\mathbf{x}_2) = \lambda_2(\mathbf{x}_1\cdot\mathbf{x}_2).$$

Por tanto, $(\lambda_1 - \lambda_2)(\mathbf{x}_1\cdot\mathbf{x}_2) = 0$. Como se ha supuesto que λ_1 y λ_2 no son iguales, debemos tener que $\mathbf{x}_1\cdot\mathbf{x}_2 = 0$.

En la Sec. 7 del Cap. 5 se dará una versión más general del Teorema 9.5 y se le aplicará a un tipo distinto de ejemplos.

Ejemplo 6. La función g de $\mathcal{R}^2$ en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

tiene valores propios que se pueden obtener de la Ecuación 9.3, es decir, resolviendo la ecuación

$$\det \begin{pmatrix} (1-\lambda) & 2 \\ 2 & (1-\lambda) \end{pmatrix} = 0,$$

es decir, $(1-\lambda)^2 - 4 = 0$. Las soluciones son $\lambda = 3$ y $\lambda = -1$. (Nótese que la transformación g no es la misma que la del Ejemplo 1, a pesar de que tiene los mismos valores propios.) La Ec. 9.4 para los vectores propios tiene dos interpretaciones, dependiendo del valor propio usado. Tenemos que

$$\lambda = 3: \quad \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\lambda = -1: \quad \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Estas ecuaciones se reducen a

$$\lambda = 3: \qquad -2x + 2y = 0,$$
$$\lambda = -1: \qquad 2x + y = 0.$$

Encontramos las soluciones

$$\lambda = 3: \qquad (x, y) = (1, 1),$$
$$\lambda = -1: \qquad (x, y) = (1, -1).$$

Los vectores $\mathbf{x}_1 = (1, 1)$, $\mathbf{x}_2 = (1, -1)$ son evidentemente ortogonales y se pueden normalizar obteniendo $\mathbf{u}_1 = (1/\sqrt{2},\ 1/\sqrt{2})$, $\mathbf{u}_2 = (1/\sqrt{2},\ -1/\sqrt{2})$. Es claro que estos vectores normalizados también son vectores propios.
Ahora podemos usar el hecho de que $\mathbf{u}_1$ y $\mathbf{u}_2$ son vectores propios que forman un conjunto ortonormal. Este último hecho nos permite expresar los vectores de $\mathcal{R}^2$ como combinaciones lineales de $\mathbf{u}_1$ y $\mathbf{u}_2$ en una forma muy sencilla. En la sección anterior vimos que

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x} \cdot \mathbf{u}_2)\mathbf{u}_2$$

para todo $\mathbf{x}$ en $\mathcal{R}^2$. Ahora, como $\mathbf{u}_1$ y $\mathbf{u}_2$ son vectores propios de la función g, podemos escribir

$$g(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{u}_1)g(\mathbf{u}_1) + (\mathbf{x} \cdot \mathbf{u}_2)g(\mathbf{u}_2).$$

Pero $g(\mathbf{u}_1) = 3\mathbf{u}_1$ y $g(\mathbf{u}_2) = -\mathbf{u}_2$, de modo que

$$g(\mathbf{x}) = 3(\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{x} \cdot \mathbf{u}_2)\mathbf{u}_2.$$

Por tanto, la acción de g tiene una descripción geométrica como la que diéramos para f en el Ejemplo 2. La única diferencia es que la expansión y la reflexión se hacen a lo largo de rectas distintas: perpendiculares en el caso de g, no perpendiculares en el caso de f.

EJERCICIOS

1. La función lineal f de $\mathcal{R}^2$ en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix}$$

tiene valores propios 7 y -5. ¿Cuáles de los vectores que siguen son vectores propios? Para los que lo son, ¿cuáles son los valores propios correspondientes?

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -4 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

2. Encuentre todos los valores propios de cada una de las funciones lineales definidas por las matrices

(a) $\begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$.

3. Encuentre un vector propio para cada valor propio de las matrices del Ejercicio 2.

4. Demuestre que 0 es un valor propio de una función lineal f si, y solo si, f no es inyectiva.

5. Demuestre que si f es una función lineal inyectiva con un valor propio λ, entonces la inversa, f^{-1}, de f tiene un autovalor $1/\lambda$.

6. Sea f una función lineal con valor propio λ.

(a) Demuestre que λ^2 es un valor propio de $f \circ f$.
(b) Demuestre que λ^n es un valor propio de la función obtenida componiendo f n veces consigo misma.

7. Sea $C^{(\infty)}(\mathcal{R})$ el espacio vectorial de las funciones infinitamente diferenciables $f(x)$ para x en $\mathcal{R}$. Entonces el operador diferencial D^2 actúa linealmente de $C^{(\infty)}(\mathcal{R})$ en $C^{(\infty)}(\mathcal{R})$.

(a) Demuestre que para todo número real λ las funciones $\cos \lambda x$ y $\operatorname{sen} \lambda x$ son vectores propios de D^2 con valor propio $-\lambda$.
(b) Sea $C^{(\infty)}[0, \pi]$ el subespacio de $C^{(\infty)}(\mathcal{R})$ formado por las funciones f tales que $f(0) = f(\pi) = 0$. Demuestre que si D^2 es restringido a $C^{(\infty)}[0, \pi]$, entonces sus únicas funciones propias son de la forma $\operatorname{sen} kx$, que corresponden a $\lambda = -k^2$, donde k es un entero.

8. Sea h una función lineal de $\mathcal{R}^2$ en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

(a) Demuestre que los vectores propios de h generan un subespacio de dimensión 1 de $\mathcal{R}^2$. Por tanto, no se puede analizar geométricamente la acción de h considerando solo su acción sobre los vectores propios.
(b) La función h se llama una **transformación deslizante.** Dé una descripción geométrica de la acción de h en $\mathcal{R}^2$.

9. (a) Encuentre los valores propios y los vectores propios correspondientes de la función lineal f de $\mathcal{R}^2$ en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

(b) Demuestre que los vectores propios de la función f de la parte (a) forman una base ortogonal para $\mathcal{R}^2$ y use este hecho para dar una descripción geométrica de la acción de f en $\mathcal{R}^2$.

(c) Generalice los resultados obtenidos en (a) y (b) a una función lineal f de $\mathcal{R}^n$ en $\mathcal{R}^n$ con matriz de la forma **diag** $(a_1, a_2, \ldots, a_n)$.

10. Encuentre los valores propios de la función g en $\mathcal{R}^2$ con matriz

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Demuestre que los vectores propios correspondientes generan $\mathcal{R}^2$ y describa la acción de g.

11. Sea $\mathcal{C}^2$ el espacio vectorial complejo de los pares $\mathbf{z} = (z_1, z_2)$ de números complejos y suponga que $\mathcal{C}^2$ tiene el producto interior complejo.

$$\mathbf{z} \cdot \mathbf{w} = z_1\bar{w}_1 + z_2\bar{w}_2.$$

(a) Demuestre que la función lineal f definida en $\mathcal{C}^2$ y con matriz

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

tiene los siguientes vectores propios y valores propios correspondientes

$$\lambda = e^{i\theta}: \quad \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

$$\lambda = e^{-i\theta}: \quad \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

(b) Demuestre que si $\operatorname{sen}\theta = 0$, entonces todo vector distinto de cero en $\mathcal{C}^2$ es un vector propio de la función f de la parte (a).
(c) Demuestre que los vectores propios de la parte (a) forman una base de $\mathcal{C}^2$.
(d) Demuestre que los vectores propios de la parte (a) son ortogonales con respecto del producto interior complejo en $\mathcal{C}^2$.

12. Explique en términos geométricos por qué una rotación alrededor del origen en $\mathcal{R}^2$ en un ángulo θ, $0 \leq \theta < 2\pi$, no tiene vectores propios salvo si $\theta = 0$ o $\theta = \pi$.

13. Suponga que $\mathcal{W}$ es un espacio vectorial complejo con un producto interior complejo tal que $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$. Entonces se dice que una función lineal f de $\mathcal{W}$ en $\mathcal{W}$ es **simétrica hermítica** si $\langle f(\mathbf{z}), \mathbf{w} \rangle = \langle \mathbf{z}, f(\mathbf{w}) \rangle$ para $\mathbf{z}$, $\mathbf{w}$ en $\mathcal{W}$.

(a) Demuestre que si f es simétrica hermítica y λ es un valor propio, entonces λ debe ser real.
(b) Demuestre que si $\mathcal{W}$ tiene dimensión compleja 2 y f está dada por la matriz compleja

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix},$$

entonces f es simétrica hermítica si, y solo si, $\beta = \bar{\gamma}$.

14. Si

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

define una función de la variable real t y que toma valores en $\mathcal{R}^2$, entonces definimos la derivada de $\mathbf{x}$ por

$$\mathbf{x}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}.$$

(a) Sea

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Demuestre que la ecuación vectorial

$$\mathbf{x}' = A\mathbf{x}$$

es lo mismo que el sistema de ecuaciones diferenciales de primer orden.

$$x' = ax + cy$$

$$y' = bx + dy$$

(b) Demuestre que $\mathbf{x}(t) = e^{\lambda t}\mathbf{c}$, donde $\mathbf{c}$ es un vector constante, es una solución de la ecuación vectorial $\mathbf{x}' = A\mathbf{x}$ si $\mathbf{c}$ es un vector propio de la matriz A con valor propio correspondiente λ.
(c) Si la matriz A es

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

encuentre las soluciones de la ecuación diferencial vectorial $\mathbf{x}' = A\mathbf{x}$ en la forma

$$\mathbf{x} = e^{\lambda_1 t}\mathbf{c}_1 + e^{\lambda_2 t}\mathbf{c}_2.$$

[*Nota.* Los vectores propios $\mathbf{c}_1$ y $\mathbf{c}_2$ están determinados salvo un múltiplo numérico distinto de cero.]

(d) Demuestre que la ecuación diferencial de segundo orden

$$(D - r_1)(D - r_2)y = 0$$

es equivalente al sistema de ecuaciones de primer orden

$$x' = r_1 x$$

$$y' = x + r_2 x,$$

donde $x = (D - r_2)y$. Encuentre la matriz A que puede ser usada para expresar este sistema en la forma de la parte (a).
(e) Demuestre que la ecuación característica de la matriz A encontrada en la parte (d) es igual a la ecuación característica del operador $(D - r_1)(D - r_2)$, tal como se la definiera en la Sec. 6.

15. (a) Demuestre que las funciones $\cos kx$ y $\operatorname{sen} kx$ son vectores propios del operador diferencial D^2 que actúa en $C^{(\infty)}$. ¿Cuáles son los valores propios correspondientes?
(b) Si restringimos el operador D^2 a las funciones reales definidas para $-\pi \leq x \leq \pi$, podemos definir un producto interno por

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

Demuestre que los vectores propios cos kx y sen lx son ortogonales con respecto de este producto interno para $k, l = 1, 2, 3, \ldots$ Use el Teorema 8.1 para deducir que los vectores propios son linealmente independientes.

(c) Demuestre que las funciones e^{ikx} y e^{-ikx} son vectores propios del operador diferencial D^2 que actúa sobre las funciones complejas. ¿Cuáles son los valores propios correspondientes?

(d) Si se usa el producto interno complejo

$$\langle f, g\rangle_* = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx,$$

demuestre que las funciones e^{ikx} son ortogonales para $k = 0, \pm 1, \pm 2, \ldots$ ¿Se puede deducir del Teorema 8.1 que los exponenciales complejos son linealmente independientes?

10. Coordenadas

En esta sección demostramos que todo problema sobre espacios vectoriales de dimensión finita y funciones lineales definidas entre ellos, puede ser reducida a un problema sobre $\mathcal{R}^n$ y matrices. Esto se hace introduciendo coordenadas en los espacios vectoriales. Las coordenadas usuales de la geometría en 2 y 3 dimensiones pueden ser consideradas como casos especiales de lo que vamos a describir. El siguiente teorema es fundamental.

10.1 Teorema

Sea $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ una base para el espacio vectorial $\mathcal{U}$, de modo que $\dim(\mathcal{U}) = n$. Para cada vector $\mathbf{x}$ en $\mathcal{U}$ existe una, y solo una, n-pla de números $r_1, \ldots, r_n$ tal que $\mathbf{x} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$.

Demostración. Como los $\mathbf{v}$ generan $\mathcal{U}$, $\mathbf{x}$ es una combinación lineal de ellos, por tanto, existe por lo menos una n-pla $(r_1, \ldots, r_n)$ que satisface la condición. Si $(s_1, \ldots, s_n)$ es otra n-pla, tal que $s_1\mathbf{v}_1 + \ldots + s_n\mathbf{v}_n = \mathbf{x} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$, entonces $(r_1 - s_1)\mathbf{v}_1 + \ldots + (r_n - s_n)\mathbf{v}_n = 0$. Como $\mathbf{v}$ son independientes, los números $r_i - s_i$ deben ser todos cero, de modo que $(s_1, \ldots, s_n)$ es igual a $(r_1, \ldots, r_n)$.

Sea $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ una base del espacio vectorial $\mathcal{U}$ y sea $\mathbf{x}$ un vector en $\mathcal{U}$. La n-pla única $(r_1, \ldots, r_n)$ tal que $\mathbf{x} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$ se llama la n-pla de **coordenadas de $\mathbf{x}$ con respecto a la base V**.

El vector de $\mathcal{R}^n$ con coeficientes $r_1, \ldots, r_n$ se llama el **vector coordenado** de $\mathbf{x}$ con respecto a la base.

Ejemplo 1. (a) Los vectores $\mathbf{v}_1 = (-1, 1, 0)$, $\mathbf{v}_2 = (0, -1, 1)$ constituyen una base para el subespacio $\mathcal{U}$ de $\mathcal{R}^3$ formado por todos los vectores cuya suma de coeficientes es 0. El vector $\mathbf{x} = (1, -3, 2)$ está en $\mathcal{U}$. Un cálculo sencillo muestra que $\mathbf{x} = -\mathbf{v}_1 + 2\mathbf{v}_2$, de modo que las coordenadas de $\mathbf{x}$ con respecto a la base $\mathbf{v}_1, \mathbf{v}_2$ son $(-1, 2)$.

(b) En el plano, sea $\mathbf{v}_1$ el vector de largo unidad a lo largo del eje horizontal y $\mathbf{v}_2$ el vector de largo unidad que forma un ángulo de 60° en el sentido contrario al de los punteros del reloj a partir de $\mathbf{v}_1$. Sea $\mathbf{x}$ el vector de largo unidad 60° más allá de $\mathbf{v}_2$, mostrado en la Fig. 16.

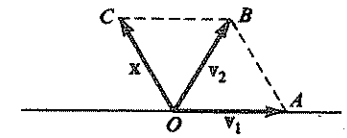


Figura 16

Usando argumentos geométricos, se ve que CB es paralelo a OA y OC es paralelo a AB (¿por qué?), de modo que $OABC$ es un paralelogramo y $\mathbf{v}_2 = \mathbf{x} + \mathbf{v}_1$. Por tanto, $\mathbf{x} = -\mathbf{v}_1 + \mathbf{v}_2$ y las coordenadas de $\mathbf{x}$ con respecto a esta base son $(-1, 1)$. Las coordenadas son una manera de representar los vectores en forma concreta, de gran utilidad para los cálculos. También son una manera concreta de representar funciones lineales. El hecho fundamental es que si se conoce el valor de una función lineal para los vectores de una base, entonces se conoce totalmente la función. Esto se debe a que un vector cualquiera $\mathbf{x}$ puede ser expresado como una combinación lineal $r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$ de los vectores de la base. Entonces, como f es lineal, $f(\mathbf{x})$ es la combinación lineal $r_1 f(\mathbf{v}_1) + \ldots + r_n f(\mathbf{v}_n)$ de los vectores $f(\mathbf{v}_i)$, con los mismos coeficientes. Sea $\mathcal{U} \xrightarrow{f} \mathcal{W}$ una función lineal y $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $W = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ bases de $\mathcal{U}$ y $\mathcal{W}$, respectivamente. No se excluye la posibilidad que ambos espacios sean iguales. **La matriz de f con respecto a las bases V y W** es, por definición, la matriz m-por-n cuya columna j es el vector coordenado de $f(\mathbf{v}_j)$ con respecto a la base W. Así, si denotamos por $A = (a_{ij})$ la matriz de f con respecto a V y W, entonces los elementos en la columna j de A son los coeficientes a_{ij} de la ecuación

$$f(\mathbf{v}_j) = a_{ij}\mathbf{w}_1 + \ldots + a_{mj}\mathbf{w}_m.$$

La matriz A contiene toda la información necesaria para evaluar $f(\mathbf{x})$ para cualquier $\mathbf{x}$. En efecto, si $\mathbf{x}$ se representa en la forma

$$\mathbf{x} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$$

con vector coordenada $\mathbf{r} = (r_1, \ldots, r_n)$, entonces

$$\begin{aligned} f(\mathbf{x}) &= \sum_{j=1}^{n} r_j f(\mathbf{v}_j) \\ &= \sum_{j=1}^{n} r_j \sum_{i=1}^{m} a_{ij}\mathbf{w}_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}r_j\right)\mathbf{w}_i. \end{aligned}$$

Podemos reconocer los coeficientes en la representación de $f(\mathbf{x})$ con respecto

de W como los elementos del producto $A\mathbf{r}$ de la matriz A y el vector coordenada $\mathbf{r}$. Notamos que si $\mathcal{U}$ y $\mathcal{W}$ son el mismo espacio, generalmente es conveniente, a pesar de no ser lógicamente necesario, tomar las bases V y $\mathcal{U}$ iguales. Hacemos esto en todos los ejemplos en los cuales $\mathcal{U}$ y $\mathcal{W}$ son iguales.

Ejemplo 2. (a) Si f es la función lineal de $\mathcal{R}^n$ en $\mathcal{R}^m$ definida por la multiplicación por una matriz A, entonces la matriz de f con respecto a las bases naturales de $\mathcal{R}^n$ y $\mathcal{R}^m$ es A. Esto no es más que el Teorema 4.2 del Cap. 1.

(b) La función lineal de $\mathcal{R}^3$ en $\mathcal{R}^3$ definida por

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix}$$

lleva el subespacio $\mathcal{U}$ generado por $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$ en sí mismo y puede ser considerada como una función lineal de $\mathcal{U}$ en $\mathcal{U}$. Tenemos

$$f(\mathbf{v}_1) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = -1\mathbf{v}_1 - 1\mathbf{v}_2$$

y $f(\mathbf{v}_2) = 1\mathbf{v}_1 + 0\mathbf{v}_2$, de modo que la matriz de f con respecto a la base $(\mathbf{v}_1, \mathbf{v}_2)$ es

$$\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}.$$

(c) Sea $\mathbf{v}_1, \mathbf{v}_2$ la base de $\mathcal{R}^2$ dada en el Ejemplo 1(b) y sea f la rotación en 60° en el sentido opuesto al de los punteros del reloj. Entonces $f(\mathbf{v}_1) = \mathbf{v}_2$ y $f(\mathbf{v}_2)$ es el vector $\mathbf{x}$ de la Fig. 16 que es igual a $-\mathbf{v}_1 + \mathbf{v}_2$. La matriz de f con respecto a la base $\mathbf{v}_1, \mathbf{v}_2$ es, por tanto,

$$\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

(d) Sea P_2 el espacio vectorial de todos los polinomios cuadráticos y $\mathcal{W} = \mathcal{R}^4$. Definamos $f(p)$, donde p es un polinomio en P_2, como el vector de $\mathcal{R}^4$ cuyas entradas son $p(1)$, $p(-1)$, $p(2)$, $p(3)$. Tomemos los polinomios 1, x, x^2 como una base de P_2 y la base natural para $\mathcal{R}^4$. Entonces, f transforma al polinomio 1 en el 4-plo $(1, 1, 1, 1)$, el polinomio x en el 4-plo $(1, -1, 2, 3)$ y x^2 en $(1, 1, 4, 9)$. Por tanto, su matriz con respecto a las bases dadas es

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}.$$

Para una base $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ de $\mathcal{U}$ podemos definir una función c_V de $\mathcal{U}$ en $\mathcal{R}^n$ haciendo $c_V(\mathbf{x})$ igual al vector coordenado de $\mathbf{x}$ con respecto a V para todo $\mathbf{x}$ en $\mathcal{U}$. La función c_V se llama la **función coordenada para la base** V. La imagen de c_V es, claramente, todo $\mathcal{R}^n$, puesto que para cualquier n-pla $(r_1, \ldots, r_n)$, tomando $\mathbf{x} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$, entonces $c_V(\mathbf{x}) = (r_1, \ldots, r_n)$. Por el Teorema 10.1, c_V es inyectiva. Por tanto, existe una función inversa c_V^{-1} de $\mathcal{R}^n$ en $\mathcal{U}$. Se comprueba fácilmente que $c_V^{-1}(r_1, \ldots, r_n) = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$. Un hecho obvio, pero importante, es el siguiente:

10.2 Teorema

Las funciones coordenadas y sus inversos, son funciones lineales.

Demostración. Para demostrar que c_V es lineal, debemos verificar que $c_V(r\mathbf{x} + s\mathbf{y}) = rc_V(\mathbf{x}) + sc_V(\mathbf{y})$. Esto sigue inmediatamente, ya que si $\mathbf{x} = \sum_{i=1}^{n} a_i\mathbf{v}_i$, y $\mathbf{y} = \sum_{i=1}^{n} b_i\mathbf{v}_i$, entonces $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^{n} (ra_i + sb_i)\mathbf{v}_i$. La linealidad de c_V^{-1} se desprende de las mismas ecuaciones.

Nótese que si $\mathcal{U} = \mathcal{R}^n$, la función coordenada c para la base natural es la función identidad $c(\mathbf{x}) = \mathbf{x}$.

Una función lineal $\mathcal{U} \xrightarrow{f} \mathcal{W}$ inyectiva y tal que su imagen es todo $\mathcal{W}$ se llama un **isomorfismo** y dos espacios tales que existe un isomorfismo entre ellos, se llaman **isomorfos**. Hemos visto que las funciones coordenadas son isomorfismos, lo cual demuestra

10.3 Teorema

Todo espacio vectorial n-dimensional es isomorfo a $\mathcal{R}^n$.

La importancia de estos conceptos está en el hecho de que dos espacios isomorfos son similares en todos sus aspectos, si se los considera como espacios vectoriales abstractos. Cualquier afirmación verdadera para un espacio (siempre que se trate de una afirmación que puede ser expresada en términos de la suma y la multiplicación numérica) se puede transformar, mediante el isomorfismo, en una afirmación verdadera para el otro espacio. Por ejemplo, si $\mathcal{U} \xrightarrow{f} \mathcal{W}$ es un isomorfismo, entonces la ecuación $\mathbf{y} = \sum_{i=1}^{n} a_i\mathbf{x}_i$ es válida en $\mathcal{U}$ si, y solo si, $f(\mathbf{y}) = \sum_{i=1}^{n} a_i f(\mathbf{x}_i)$ en $\mathcal{W}$.

Usando funciones coordenadas, podemos dar otra descripción de la matriz de una función lineal. Dada $\mathcal{U} \xrightarrow{f} \mathcal{W}$ y bases $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ de $\mathcal{U}$ y $W = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ de $\mathcal{W}$, la función $c_W \circ f \circ c_V^{-1}$ es una función de $\mathcal{R}^n$ en $\mathcal{R}^m$. Por tanto, ella está dada por la multiplicación por una matriz A, afirmamos que

A es precisamente la matriz de f con respecto a las bases V y W. Para demostrarlo, notemos que la columna j de A es $(c_W \circ f \circ c_V^{-1})(\mathbf{e}_j) = c_W(f(\mathbf{v}_j))$, ya que $c_V^{-1}(\mathbf{e}_j) = \mathbf{v}_j$. Ahora bien, $c_W(f(\mathbf{v}_j))$ es (por definición de c_W) el vector coordenado de $f(\mathbf{v}_j)$ con respecto a la base W, y ésta es precisamente la forma como se definió la matriz de f con respecto a las bases V y W. Esta descripción facilita la demostración de la siguiente generalización del Teorema 4.4 del Cap. 1. En términos generales, se afirma que la multiplicación de matrices corresponde a la composición de funciones, siempre que se usen bases en forma consistente.

10.4 Teorema

Sean f y g funciones lineales tales que $\mathfrak{U} \xrightarrow{f} \mathfrak{V}$ y $\mathfrak{V} \xrightarrow{g} \mathfrak{W}$. Sean U, V y W bases para $\mathfrak{U}$, $\mathfrak{V}$ y $\mathfrak{W}$; A la matriz de f con respecto a las bases U y V y B la matriz de g con respecto a las bases V y W. Entonces la matriz de $g \circ f$ con respecto a las bases U y W es BA.

Demostración. Por el Teorema 4.3 del Cap. 1 y la caracterización de las matrices de funciones usando funciones coordenadas, esto significa simplemente que

$$(c_W \circ g \circ c_V^{-1}) \circ (c_V \circ f \circ c_U^{-1}) = c_W \circ (g \circ f) \circ c_U^{-1},$$

lo cual es claro, ya que la composición de funciones es asociativa.

El caso especial en que $\mathfrak{U}$, $\mathfrak{V}$ y $\mathfrak{W}$ son iguales y se considera una sola base particularmente importante. Si $\mathfrak{V} \xrightarrow{f} \mathfrak{V}$ tiene matriz A, entonces $f \circ f$ tiene matriz A^2, $f \circ f \circ f$ tiene matriz A^3, etc.

Ejemplo 3. (a) Para la función f del Ejemplo 2(b), es claro que $f \circ f \circ f$ es la identidad, y es fácil verificar que

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}^3 = I.$$

(b) La diferenciación es una función lineal del espacio de polinomios en sí mismo. Si denotamos p' por $D(p)$ para los elementos del espacio de los polinomios cuadráticos y usamos la base $(1, x, x^2)$, como en el Ejemplo 2(d), entonces la matriz de D es

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es fácil comprobar que

$$A^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

y $A^3 = 0$, lo cual corresponde al hecho de que la tercera derivada de un polinomio cuadrático es 0.

La matriz de una función lineal con respecto a un par de bases depende, por supuesto, de las bases. Como la función queda determinada completamente por su matriz con respecto a un par de bases cualquiera, debe ser posible expresar la matriz de una función con respecto a un par de bases en términos de la matriz con respecto a otro par.

Veamos un ejemplo. Supongamos que f tiene la matriz

$$\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \end{pmatrix}$$

con respecto a las bases naturales de $\mathfrak{R}^2$ y $\mathfrak{R}^3$. Consideremos las bases

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

en $\mathfrak{R}^3$ y

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

en $\mathfrak{R}^2$. Para encontrar la matriz de f con respecto a $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ y $(\mathbf{w}_1, \mathbf{w}_2)$, calculamos

$$f(\mathbf{v}_1) = \begin{pmatrix} 5 \\ 10 \end{pmatrix} = 5\mathbf{w}_1,$$

$$f(\mathbf{v}_2) = \begin{pmatrix} 2 \\ 7 \end{pmatrix} = 8\mathbf{w}_1 - 3\mathbf{w}_2, \qquad f(\mathbf{v}_3) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -4\mathbf{w}_1 + 3\mathbf{w}_2.$$

Por tanto, la matriz buscada es

$$\begin{pmatrix} 5 & 8 & -4 \\ 0 & -3 & 3 \end{pmatrix}.$$

Una base $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ puede ser descrita en términos de otra base $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ mediante la matriz cuya columna j es el vector coordenado $\mathbf{v}_j$ con respecto a X. Esta matriz define una función de $\mathfrak{R}^n$ en $\mathfrak{R}^n$, que, como puede verificarse mirando su efecto en los vectores básicos $\mathbf{e}_j$, es simplemente $c_X \circ c_V^{-1}$. La multiplicación por esta matriz transforma las coordenadas de un vector $\mathbf{w}$ con respecto a V en las coordenadas con respecto a X, ya que $c_X(\mathbf{w}) = (c_X \circ c_V^{-1}) \circ c_V(\mathbf{w})$. La matriz inversa expresa las $\mathbf{x}$ en términos de los $\mathbf{v}$ y corresponde a $c_V \circ c_X^{-1}$. En el ejemplo del párrafo anterior, las matrices que expresan V y W en términos de las bases naturales de $\mathfrak{R}^2$ y $\mathfrak{R}^3$ son

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

Como $\mathbf{e}_1 = -3\mathbf{w}_1 + 2\mathbf{w}_2$ y $\mathbf{e}_2 = 2\mathbf{w}_1 - \mathbf{w}_3$, la matriz que expresa las e en términos de las w es

$$\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix},$$

que, como se verifica fácilmente, es

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1}.$$

En la situación general, tenemos un espacio $\mathcal{U}$ con bases $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ y $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ y un espacio $\mathcal{W}$ con bases $Y = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ y $W = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$.

Sean P y Q las matrices que expresan V en términos de X y W en términos de Y, y sea f una función lineal de $\mathcal{U}$ en $\mathcal{W}$ cuya matriz con respecto a X e Y es A. Entonces la matriz de f con respecto a V y W es $Q^{-1}AP$. La manera más fácil de comprobar esto es usando funciones coordenadas. La matriz de f con respecto a V y W corresponde a

$$c_W \circ f \circ c_V^{-1} = (c_W \circ c_Y^{-1}) \circ (c_Y \circ f \circ c_X^{-1}) \circ (c_X \circ c_V^{-1}),$$

y los tres factores de la derecha corresponden a Q^{-1}, A y P, respectivamente. En el ejemplo anterior,

$$Q^{-1}AP = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 8 & -4 \\ 0 & -3 & 3 \end{pmatrix},$$

que es el mismo resultado obtenido anteriormente.

Si $\mathcal{U}$ y $\mathcal{W}$ son iguales y X e Y son la misma base como también lo son V y W, entonces P y Q son iguales y la nueva matriz para f es $P^{-1}AP$. Dos matrices A y B son **similares** si existe una matriz invertible P tal que $B = P^{-1}AP$. Algunas propiedades de la similaridad son dadas en los ejercicios.

Ejemplo 4. En el espacio de los polinomios cuadráticos la función derivada D tiene matriz

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

con respecto a la base $(1, x, x^2)$ [véase Ejemplo 3(b)]. Los polinomios $1 + x$, $x + x^2$, $1 + x^2$ también forman una base, dada en términos de 1, x, x^2 por la matriz

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

La matriz inversa es

$$\tfrac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

Con respecto a esta nueva base, D tiene la matriz

$$\tfrac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$= \tfrac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \tfrac{1}{2}\begin{pmatrix} 1 & 3 & 2 \\ -1 & 1 & 2 \\ 1 & -1 & -2 \end{pmatrix}$$

EJERCICIOS

1. Sean $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, y $\mathbf{v}_3 = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}$. Verifique que $V = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ es una base de $\mathbb{R}^3$ y encuentre los vectores coordenados de $\mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, y $\mathbf{x}_3 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$ con respecto de V.

2. Encuentre la matriz de la rotación de $\mathcal{R}^2$ en 45° con respecto de

(a) La base $\mathbf{x}_1 = \begin{pmatrix}1\\1\end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix}0\\-2\end{pmatrix}$. $\left[\textit{Resp}: \begin{pmatrix}0 & \sqrt{2}\\ -1/\sqrt{2} & \sqrt{2}\end{pmatrix}.\right]$

(b) La base de los vectores unitarios de las direcciones de $\mathbf{x}_1$ y $\mathbf{x}_2$.

3. Demuestre que las matrices elementales E_{ij} de forma m-por-n constituyen una base para el espacio vectorial de las matrices m-por-n. ¿Cuál es la dirección de este espacio?

4. Sea $A = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$. Demuestre que cada una de las siguientes funciones del espacio de las matrices 2-por-2 en sí mismo es lineal. y encuentre la matriz de cada una de ellas con respecto a la base del Ejercicio 3.

(a) $f(X) = AX$. (b) $g(X) = AXA^{-1}$.

5. Encuentre bases para los núcleos de las funciones definidas en el espacio de las matrices 2-por-2, dadas por

(a) $f(X) = AX - XA$, (b) $g(X) = AX - X^tA^t$,

donde A es la matriz del Ejercicio 4.

6. Sea $\mathcal{U} \xrightarrow{f} \mathcal{W}$ un isomorfismo. Demuestre que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ es una base de $\mathcal{U}$ si, y solo si, $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ es una base de $\mathcal{W}$. Esto demuestra que dos espacios isomorfos tienen la misma dimensión.

7. Para dos espacios vectoriales $\mathcal{U}$ y $\mathcal{W}$ cualesquiera, sea $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ el conjunto de todas las funciones lineales de $\mathcal{U}$ en $\mathcal{W}$. Si f y g están en $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ y r es un número, definimos funciones

$$f + g \quad \text{por} \quad (f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$$

y

$$r \cdot f \quad \text{por} \quad (r \cdot f)(\mathbf{v}) = r \cdot f(\mathbf{v}).$$

Demuestre que $f + g$ y $r \cdot f$ son lineales; es decir, que están en $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ y que, con estas operaciones, $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ es un espacio vectorial.

8. (a) Demuestre que $\mathfrak{L}(\mathcal{R}^n, \mathcal{R}^m)$ es isomorfo al espacio de las matrices m-por-n.
(b) Demuestre que si dim $(\mathcal{U}) = n$ y dim $(\mathcal{W}) = m$ y si se eligen bases en $\mathcal{U}$ y $\mathcal{W}$, entonces, haciendo corresponder a cada f en $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ su matriz con respecto a esas bases, se obtiene un isomorfismo entre $\mathfrak{L}(\mathcal{U}, \mathcal{W})$ y el espacio de las matrices m-por-n.

9. Sea P_n el espacio de los polinomios de grado menor o igual que n. Demuestre que la función t definida por $t(p(x)) = p(x + 1)$ (de modo que, por ejemplo, $t(2x + 1) = 2x + 3$) es una función lineal de P_n en sí mismo.

(a) Escriba la matriz de t con respecto a la base 1, x, x^2 para P_2.
(b) Verifique mediante el cálculo matricial que, en P_2,

$$t = 1 + D + \frac{D^2}{2},$$

donde D es la función derivada del Ejemplo 3(b). («1» debe interpretarse como la función identidad y D^2 como $D \circ D$.)

(c) Demuestre que en P_n

$$t = 1 + D + \frac{D^2}{2} + \frac{D^3}{6} + \ldots + \frac{D^n}{n!}$$

[*Indicación*: Use directamente las definiciones, sin usar coordenadas.]

10. Demuestre que si A y B son similares, entonces

(a) $\det A = \det B$.
(b) A^{-1} es similar a B^{-1} (si los inversos existen).
(c) Para cualquier entero positivo, A^n es similar a B^n.

11. La **traza** de una matriz cuadrada es, por definición, la suma de los elementos de su diagonal y se denota por tr (A). Por ejemplo, $\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} = 1 + 4 = 5$, y si I es la matriz identidad de forma n-por-n, entonces tr $(I) = n$.

(a) Demuestre que si A y B son matrices cuadradas del mismo tamaño, entonces tr $(AB) =$ tr (A) tr (B).
(b) Demuestre que si A y B son similares, entonces tr $(A) =$ tr (B). [*Indicación*: Esto puede ser demostrado en una línea aplicando la parte (a) a un par adecuado de matrices.]

12. Sea $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}^3$ una rotación y $\mathbf{v}_1$ un vector unitario en la dirección del eje de la rotación. Si $\mathbf{v}_2$ y $\mathbf{v}_3$ se eligen de modo que $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ sea una base ortonormal, entonces $f(\mathbf{v}_1) = \mathbf{v}_1$ y $f(\mathbf{v}_2)$ y $(\mathbf{v}_3)$ son perpendiculares a $\mathbf{v}_1$.

(a) Demuestre que la matriz de f con respecto a $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ es

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\alpha & -\operatorname{sen}\alpha\\ 0 & \operatorname{sen}\alpha & \cos\alpha\end{pmatrix},$$

donde α es el ángulo de rotación. (Compare con el Ejercicio 4 de la Sec. 4 del Cap. 1.)
(b) Use el resultado del Ejercicio 11, para demostrar que si A es la matriz de una rotación en un ángulo α con respecto a *cualquier* base, entonces $\cos\alpha = \frac{1}{2}(\text{tr}\,(A) - 1)$.

13. La función lineal en $\mathcal{R}^3$ definida por

$$f(\mathbf{e}_1) = \mathbf{e}_2, \quad f(\mathbf{e}_2) = \mathbf{e}_3, \quad f(\mathbf{e}_3) = \mathbf{e}_1$$

es claramente una rotación con eje la recta $x_1 = x_2 = x_3$. Encuentre su ángulo por consideraciones geométricas (¿qué es $f \circ f \circ f$?) y compruebe el resultado del Ejercicio 12.

[*Resp.* 120°.]

14. Demuestre que la matriz

$$A = \begin{pmatrix} \frac{2}{7} & \frac{6}{7} & \frac{3}{7}\\ \frac{3}{7} & \frac{2}{7} & -\frac{6}{7}\\ -\frac{6}{7} & \frac{3}{7} & -\frac{2}{7}\end{pmatrix}$$

representa una rotación de $\mathcal{R}^3$ y encuentre su eje y su ángulo. (Si $\mathbf{x}$ es un vector unitario en la dirección del eje, entonces $A\mathbf{x} = \mathbf{x}$ o $(A - I)\mathbf{x} = 0$. Para demostrar que A representa una rotación, encuentre su matriz con respecto a una base ortonormal que incluya a $\mathbf{x}$.) [*Resp.* El eje es $(3, 3, -1)$.]

3

Funciones y derivadas

1. Funciones vectoriales

Esta sección tiene como propósito familiarizar al lector con algunos ejemplos específicos de funciones vectoriales. Estos ejemplos serán dados por fórmulas que pueden ser analizadas en términos de funciones elementales del cálculo de una variable. Cuando sea posible, daremos una representación pictórica de la función, lo cual a menudo puede hacerse en más de una forma. Más adelante consideraremos ejemplos del tipo de los considerados aquí, obteniendo así una mejor comprensión de ellos.

Ejemplo 1. La función $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}^2$, definida por

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 \\ x + y + z \end{pmatrix}.$$

tiene como dominio todo $\mathcal{R}^3$, pero su imagen contiene solo a los vectores de $\mathcal{R}^2$ cuya primera coordenada es no negativa.

Otra función $\mathcal{R}^3 \xrightarrow{g} \mathcal{R}^2$ está definida por

$$g(x, y, z) = \begin{pmatrix} 3x + 4y \\ 3y + 5z \end{pmatrix} = \begin{pmatrix} 3 & 4 & 0 \\ 0 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

y g es lineal, ya que está dada por multiplicación por una matriz. La función $\mathcal{R}^2 \xrightarrow{h} \mathcal{R}^2$, definida por

$$h(x, y) = \begin{pmatrix} \sqrt{1 - x^2 - y^2} \\ x + y \end{pmatrix},$$

para los vectores (x, y), tales que $x^2 + y^2 \leq 1$, tiene como dominio un disco circular D en $\mathcal{R}^2$. A menudo se usa la notación $f: D \xrightarrow{f} \mathcal{R}^m$ para denotar una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ cuyo dominio D está contenido en $\mathcal{R}^n$.

Ejemplo 2. Las funciones definidas por las fórmulas que siguen tienen como espacio dominio y espacio imagen al conjunto $\mathcal{R}$ de los números reales.

$$f(x) = x^3,$$
$$g(x) = \sqrt{x}, \qquad x \geq 0,$$
$$h(x) = \operatorname{sen} x.$$

El dominio y la imagen de f son iguales a todo $\mathcal{R}$. De acuerdo con las interpretaciones habituales del símbolo raíz cuadrada, la imagen de g es el conjunto de los números reales no negativos que es también el dominio de g. El dominio de h es todo $\mathcal{R}$ y su imagen es el intervalo $-1 \leq y \leq 1$. Para poner énfasis en el dominio, la función g puede ser denotada por

$$g: [0, \infty) \to \mathcal{R},$$

donde $[0, \infty)$ es el intervalo $0 \leq x < \infty$.

Un vector $\mathbf{x}$ de $\mathcal{R}^n$ cuyas coordenadas son los números reales $x_1, \ldots, x_n$, será expresado como una n-pla horizontal o como una matriz columna. Es decir, lo denotaremos por

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \text{ y por } (x_1, \ldots, x_n)$$

La razón para usar columnas está en la definición de multiplicación de matrices. Si una función está determinada por una matriz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

generalmente escribiremos

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

para denotar el valor de la función en

$$\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{o} \quad (x, y).$$

Una función, cuya imagen es un subconjunto del espacio $\mathfrak{R}$ de todos los números reales, se llama una función **real.** Toda función $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ determina un conjunto de funciones reales $f_1, \ldots, f_m$, llamadas las **funciones coordenadas** de f, ponemos $f_i(\mathbf{x})$ igual a la coordenada i de $f(\mathbf{x})$. Entonces

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$$

para todo $\mathbf{x}$ en el dominio de f, y f_i se llama la i-ésima función coordenada de f.

Ejemplo 3. Consideramos la función vectorial

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ xy + yz + zx \\ xyz \end{pmatrix}.$$

Las funciones coordenadas de f son las tres funciones reales

$$f_1(x, y, z) = x + y + z,$$
$$f_2(x, y, z) = xy + yz + zx,$$
$$f_3(x, y, z) = xyz.$$

El **grafico** de una función f es, por definición, el conjunto de todos los pares ordenados $(\mathbf{x}, f(\mathbf{x}))$ con $\mathbf{x}$ en el dominio de f. En el estudio de funciones reales de una variable real, los gráficos son de gran utilidad. Por ejemplo, el gráfico de la función definida por $f(x) = x^2 - 2$ es el conjunto de todos los pares ordenados (x, y) con $y = x^2 - 2$, es decir, el subconjunto del plano xy formado por los puntos de la parábola de la Fig. 1. Otro ejemplo del gráfico de una función $\mathfrak{R}^2 \xrightarrow{f} \mathfrak{R}$ es el subconjunto de $\mathfrak{R}^3$ formado por todos los puntos

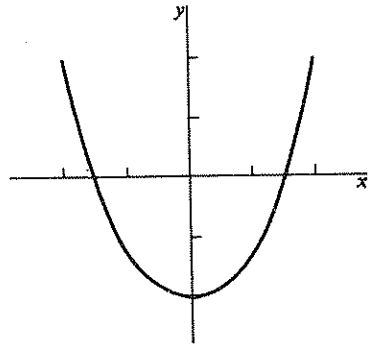


Figura 1

$$(x, y, z) = (x, y, f(x, y)).$$

La utilidad del gráfico como método para visualizar una función está limitada al caso de funciones $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ con $n + m \leq 3$.

Ejemplo 4. Para obtener el gráfico de la función f definida por

$$f(x, y) = x^2 + y^2,$$

recordemos que éste está formado por todos los puntos de $\mathfrak{R}^3$ de la forma $(x, y, z) = (x, y, x^2 + y^2)$. Una manera de encontrar el gráfico es dibujar varios de estos puntos, por ejemplo, (1, 1, 2), (0, 1, 1), (1, 0, 1), etc. Otra manera es dibujar algunas de las curvas que yacen en el gráfico. En particular, podemos dibujar las secciones del gráfico obtenidas manteniendo x o y fijo y haciendo variar el otro. El resultado es una curva en un plano paralelo al plano yz o al plano xz. La Fig. 2 muestra estas curvas. Hemos elegido $x = \frac{3}{2}$ como uno de

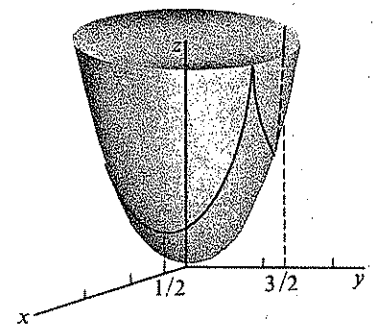


Figura 2

ellos. El resultado es el conjunto de puntos $(x, y, z) = (\frac{3}{2}, y, \frac{9}{4} + y^2)$. Para dibujarlos podemos considerar primero los puntos $(y, z) = (y, \frac{9}{4} + y^2)$ en el plano yz. Luego movemos la curva resultante $\frac{3}{2}$ unidades en la dirección positiva del eje n. Análogamente, si elegimos $y = \frac{1}{2}$, dibujamos la curva formada por los puntos $(x, y, z) = (x, \frac{1}{2}, x^2 + \frac{1}{4})$. Haciendo esto para varios valores de x e y, se obtiene una figura mejor que si solo consideramos puntos individuales.

Ejemplo 5. Consideremos la función $\mathfrak{R} \xrightarrow{f} \mathfrak{R}^2$ definida por

$$f(t) = \begin{pmatrix} \cos t \\ \operatorname{sen} t \end{pmatrix}, \quad \text{para todo } t \text{ en } \mathfrak{R}.$$

Como el largo del vector $\mathbf{x} = (x, y)$ es $|\mathbf{x}| = \sqrt{x^2 + y^2}$, tenemos que $|f(t)| = \sqrt{\cos^2 t + \operatorname{sen}^2 t} = 1$. Por tanto, la imagen de f es un subconjunto del círculo unitario $|\mathbf{x}| = 1$ de $\mathcal{R}^2$. El número t se interpreta geométricamente como el ángulo en radianes entre el vector $f(t)$ y el eje x positivo. Al variar t en $\mathcal{R}$, el círculo unitario es cubierto infinitas veces. Por tanto, la imagen de f es todo el círculo unitario. Sin embargo, el círculo no es el gráfico de f, ya que éste es un subconjunto de $\mathcal{R}^3$ formado por todos los puntos en una espiral de eje, el eje t. Véase Fig. 3(b). Lo que hemos hecho es dibujar los puntos de la forma $(t, x, y) = (t, \cos t, \operatorname{sen} t)$ que forman el gráfico de la función.

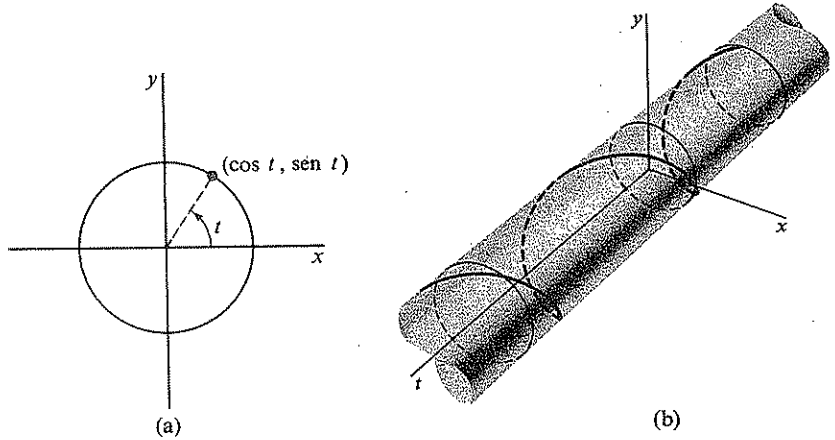


Figura 3

Ejemplo 6. Sea f la función vectorial definida por

$$f(t) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}, \qquad -\infty < t < \infty.$$

El gráfico de f es un subconjunto de $\mathcal{R}^4$, de modo que no trataremos de dibujarlo. Sin embargo, podemos dibujar su imagen. Haciendo $z = 0$, tenemos las ecuaciones $x = t$, $y = t^2$, lo cual es equivalente a $y = x^2$. Por tanto, la proyección de la imagen de f en el plano xy es el gráfico de $y = x^2$. De la misma forma, en el plano yz obtenemos $y = z^{2/3}$ y $z = x^3$ en el plano xz. Partiendo de esta información, hemos dibujado en la Fig. 4 la parte de la imagen de f que está sobre el primer cuadrante del plano xy.

Al dibujar la Fig. 4, hemos designado los ejes en la manera asociada generalmente con la orientación derecha. Si intercambiamos x e y en la designación, la figura sería, por supuesto, diferente. Con un cambio similar en la Fig. 5(b), se obtendría una espiral que gira hacia la izquierda en lugar de hacia la derecha.

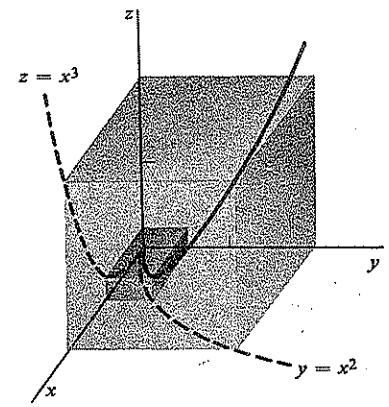


Figura 4

Para clarificar la relación entre distintas figuras, elegimos siempre la orientación derecha para los ejes 3-dimensionales. En la Sec. 7 del Cap. 1, se puede encontrar una discusión sobre el concepto de orientación.

Ejemplo 7. Consideremos la función

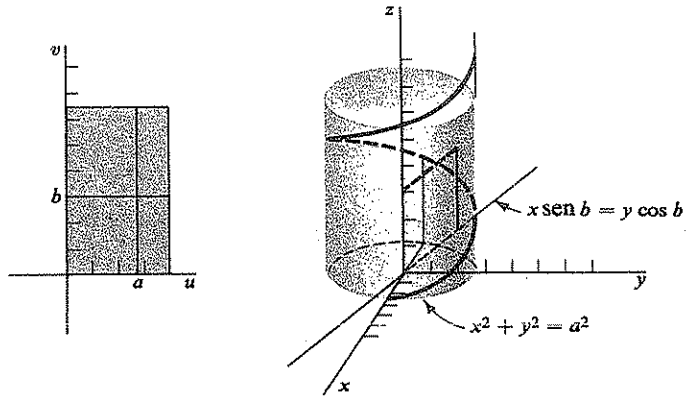
$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u\cos v \\ u\operatorname{sen} v \\ v \end{pmatrix} \qquad \begin{matrix} 0 \le u \le 4, \\ 0 \le v \le 2\pi. \end{matrix}$$

El dominio de f es el rectángulo sombreado de la Fig. 5(a). Para dibujar su imagen, procedemos como sigue: Elegimos un número a en el intervalo $0 \le u \le 4$ y hacemos $u = a$. Entonces,

$$\left.\begin{aligned} x &= a\cos v \\ y &= a\operatorname{sen} v \\ z &= v\operatorname{sen} \end{aligned}\right\} \qquad 0 \le v \le 2\pi,$$

y $x^2 + y^2 = a^2$. Interpretamos v como la distancia a lo largo del eje z y como el ángulo entre $(x, y, 0)$ y el eje x. Por tanto, la imagen por f del segmento de recta $u = a$, para $0 \le 2 \le 2\pi$ [véase Fig. 5(b)] es la espiral cuya proyección en el plano xy es el círculo de radio a y cuyo eje es el eje z [véase Fig. 5(b)]. A continuación, elegimos un número b en el intervalo $0 \le v \le 2\pi$ y hacemos $v = b$. Entonces

$$\left.\begin{aligned} x &= u\cos b \\ y &= u\operatorname{sen} b \\ z &= b \end{aligned}\right\} \qquad 0 \le u \le 4,$$



(a) (b)

Figura 5

y $x^2 + y^2 = u^2$. La imagen por f del segmento $v = b$, $0 \leq u \leq 4$ [véase Fig. 5(a)] es el segmento de recta de largo $x \operatorname{sen} b = y \cos b$, $z = b$, donde x varía entre 0 y 4 cos b [véase Fig. 5(b)]. Haciendo variar a y b, veremos que la imagen de f es la superficie espiral de la Fig. 6.

Nótese que, en los Ejemplos 5 y 6, la imagen de f es una curva, mientras que en el Ejemplo 7 es una superficie. Análogamente, el gráfico de f en el Ejemplo 4 es una superficie y en el Ejemplo 5 es una curva. Esta evidencia sugiere que cuando la dimensión del espacio dominio 1 se obtiene una curva y que, cuando ésta es 2, se obtiene una superficie. En la sección última del Cap. 4 se discutirán algunas excepciones.

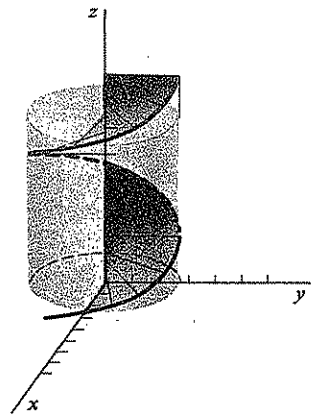


Figura 6

Las curvas y superficies de las Figs. 2, 3, 4 y 6 están relacionadas con las funciones vectoriales que las definen en dos formas esencialmente diferentes. La superficie de la Fig. 2 es el gráfico de la función que la define, como también lo es la curva espiral de la Fig 3(b). Diremos que esta superficie y curva han sido definidas **explícitamente.** Por otra parte, la curva de la Fig. 4 y la superficie de la Fig. 6 son las imágenes de ciertas funciones. Se dice que ellas han sido definidas **paramétricamente.** Esta terminología es la habitual en la discusión de funciones reales de una variable real. Por ejemplo, la función

$$f(x) = \sqrt{16 - x^2}$$

define explícitamente la mitad superior del círculo de radio 4, curva que puede ser definida paramétricamente mediante el par de funciones

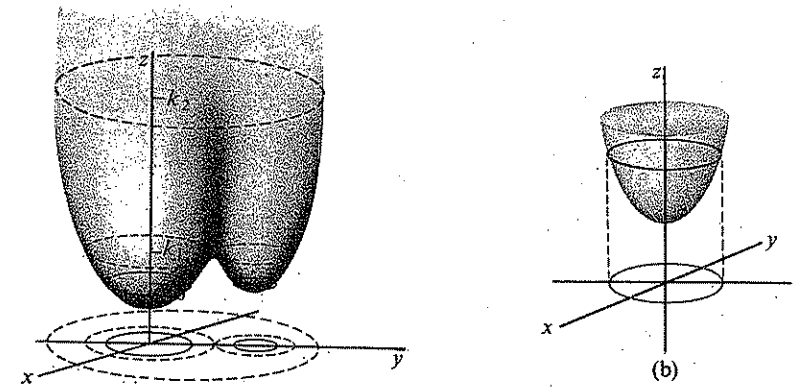
$$\begin{cases} x(t) = 4 \cos t, \\ y(t) = 4 \operatorname{sen} t, \end{cases} \qquad 0 \leq t \leq \pi.$$

Las representaciones paramétricas de rectas en el espacio tridimensional han sido estudiadas en el Cap. 1. Sean $\mathbf{x}_1$ y $\mathbf{x}_2$ dos vectores en $\mathcal{R}^3$. Si $\mathbf{x}_1 \neq 0$, entonces la imagen de la función $\mathcal{R} \xrightarrow{f} \mathcal{R}^3$ definida por

$$f(t) = t\mathbf{x}_1 + \mathbf{x}_2, \qquad -\infty < t < \infty,$$

es una recta definida paramétricamente.

Una curva o superficie también puede ser definida **implícitamente** como el conjunto de nivel de una función. Un conjunto S es un **conjunto de nivel** de f si para algún punto $\mathbf{k}$ en la imagen de f, S está formado por todos los puntos $\mathbf{x}$ en el dominio de f tales que $f(\mathbf{x}) = \mathbf{k}$. Los casos más fáciles de ser visualizados son las funciones f de $\mathcal{R}^2$ en $\mathcal{R}$. Entonces el gráfico de f puede ser dibujado como la superficie de la Fig. 7. Los conjuntos de nivel correspondientes a los



(a) (b)

Figura 7

valores k_0, k_1 y k_2 son las curvas en el plano xy definidas implícitamente por $f(x, y) = k_0$, o k_1, o k_2. A $f(x, y) = 1$ corresponde $x^2 + y^2 = 0$, lo cual determina el punto único $(x, y) = (0, 0)$. Véase la Fig. 7(b). Los conjuntos de nivel se usan generalmente en los mapas topográficos para mostrar los cambios en la elevación del terreno en intervalos regulares.

Ejemplo 8. Consideremos la función f definida por $f(x, y, z) = x^2 + y^2 + z^2$. El subconjunto S de $\mathcal{R}^3$ formado por todos los puntos (x, y, z) que satisfacen la ecuación

$$x^2 + y^2 + z^2 = k \tag{1}$$

está definido implícitamente por la Ec. (1). Si $k > 0$, obtenemos una esfera de radio $\sqrt{k}$ con centro en (0, 0, 0), puesto que $x^2 + y^2 + z^2$ es el cuadrado de la distancia de (x, y, z) a (0, 0, 0). Si $k = 0$, obtenemos el punto (0, 0, 0). Por último, si $k < 0$, el conjunto de nivel correspondiente S es vacío. En este ejemplo, el gráfico de f es un subconjunto de $\mathcal{R}^4$ y no puede ser dibujado. En todo caso, es deseable dibujar los conjuntos de nivel para obtener así una idea de que puntos del dominio son enviados por f a ciertos puntos fijos de la imagen. Algunos de estos conjuntos se muestran en la Fig. 8 y son un punto y dos esferas concéntricas.

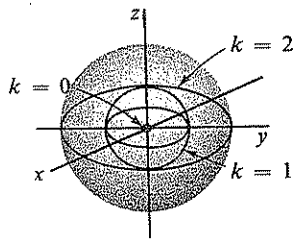


Figura 8

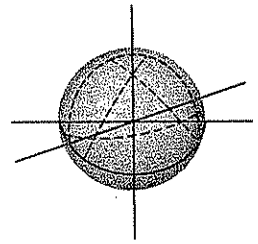


Figura 9

Ejemplo 9. Sea $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}^2$ definida por

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 \\ x + y + z \end{pmatrix}$$

Describiremos el conjunto de nivel γ definido implícitamente en $\mathcal{R}^3$ por la ecuación $f(x, y, z) = (2, 1)$. Es decir, γ está formado por todos los (x, y, z) tales que

$$\begin{aligned} x^2 + y^2 + z^2 &= 2 \\ x + y + z &= 1. \end{aligned} \tag{2}$$

Hemos visto en el Ejemplo 8 que $x^2 + y^2 + z^2 = 2$ define implícitamente una

esfera S de radio $\sqrt{2}$. La ecuación $x + y + z = 1$ define un plano P, ya que ella es satisfecha por todos los puntos (x, y, z) tales que

$$(1, 1, 1) \cdot (x, y, z - 1) = 0.$$

De hecho, P es evidentemente el plano que contiene los puntos (1, 0, 0), (0, 1, 0) y (0, 0, 1) y se lo representa en la Fig. 9.

El conjunto de nivel determinado por $f(x, y, z) = (2, 1)$ es, por tanto, el círculo C formado por todos los puntos que satisfacen *ambas* ecuaciones (2); es decir, la intersección de S y P.

Resumamos las definiciones de representaciones explícita, paramétrica e implícita. Un conjunto S está definido:

1. **explícitamente** si S es el **gráfico** en $\mathcal{R}^{n+m}$ de la función

$$\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m.$$

2. **paramétricamente** si S es la **imagen** en $\mathcal{R}^m$ de la función

$$\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m.$$

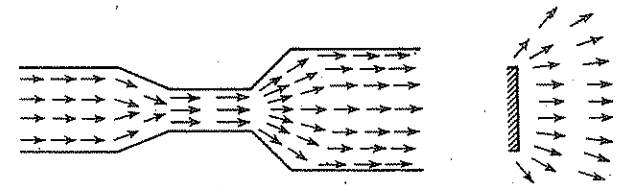
3. **implícitamente** si para alguna función

$$\mathcal{R}^{n+m} \xrightarrow{f} \mathcal{R}^m,$$

S es un **conjunto de nivel** de f; es decir, para algún punto $\mathbf{k}$ en $\mathcal{R}^m$, S es el conjunto de todos los $\mathbf{x}$ en el **dominio** de f, tales que $f(\mathbf{x}) = k$.

Un conjunto S definido en una de estas tres maneras será llamado una curva o superficie siempre que f satisfaga ciertas condiciones de suavidad que serán descritas en las últimas secciones del Cap. 4. Por ahora, usaremos estos términos informalmente.

Cuando los espacios dominio e imagen de una función vectorial son iguales, es a menudo conveniente imaginarse los vectores $\mathbf{x}$ del dominio como puntos



(a) (b)

Figura 10

y los vectores imagen $f(\mathbf{x})$ como flechas con su cola en $\mathbf{x}$. Esto se hace, por ejemplo, al representar el flujo 2-dimensional de un fluido, donde el vector imagen en cada punto es la velocidad del flujo. Véase Fig. 10(a). Otro ejemplo es un

campo eléctrico, donde el valor de la función en un punto es la fuerza que ejerce el campo sobre una carga unitaria colocada en ese punto. Véase Fig. 10(b). Cuando se considera de esta manera a una función vectorial, se la llama a veces un **campo vectorial.** Visualizamos un campo vectorial como flechas que salen de cada punto. Por tanto, el espacio dominio y el espacio imagen deben tener la misma dimensión.

Ejemplo 10. Un ejemplo sencillo de un campo vectorial es dado por la función $f(x, y) = (x, y/2)$. Para dibujarlo, ubicamos puntos (x, y) y puntos $(x, y/2)$. Entonces trasladamos la flecha dirigida de $(0, 0)$ a $(x, y/2)$ paralelamente hasta que su cola esté en el punto (x, y). Algunos de estos vectores son mostrados en la Fig. 11.

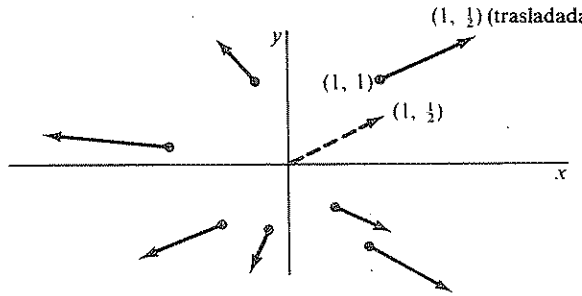


Figura 11

EJERCICIOS

1. Supongamos que la temperatura en un punto (x, y, z) del espacio está dada por $T(x, y, z) = x^2 + y^2 + z^2$. Una partícula se mueve de modo que, en el instante t su posición está dada por $(x, y, z) = (t, t^2, t^3)$. Encuentre la temperatura en el punto ocupado por la partícula en $t = \frac{1}{2}$. ¿Cuál es la velocidad de cambio de temperatura de la partícula cuando $t = \frac{1}{2}$?
2. Supongamos que la densidad por unidad de volumen de una caja cúbica de lado 2, varía directamente con la distancia al centro e inversamente con $1 + t^2$, donde t es el tiempo. Si la densidad en una de las esquinas de la caja es 1 cuando $t = 0$, encuentre una fórmula para la densidad en cualquier punto y en cualquier instante. ¿Cuál es la velocidad de cambio de la densidad en un punto a $\frac{1}{2}$ unidades del centro de la caja, en el instante $t = 1$?
3. Considere la función $f(x, y) = \sqrt{4 - x^2 - y^2}$.
   (a) Dibuje el dominio de f (tómelo lo más grande que se pueda).
   (b) Dibuje el gráfico de f.
   (c) Dibuje la imagen de f.
4. Sea $\mathbb{R} \xrightarrow{g} \mathbb{R}^2$ la función definida por

$$g(t) = \begin{pmatrix} 2\cos t \\ 3 \operatorname{sen} t \end{pmatrix}, \qquad 0 \leq t \leq 2\pi.$$

   (a) Dibuje la imagen de g.
   (b) Dibuje el gráfico de g.
5. Una transformación del plano xy en el plano uv está definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y(1 + x^2) \end{pmatrix}.$$

   ¿Cuáles son las imágenes de las rectas horizontales del plano xy? ¿Cuáles son las funciones coordenadas de f?
6. Para cada una de las siguientes funciones lineales (a) ¿Cuál es su dominio? (b) Describa y dibuje su imagen. (c) Describa y dibuje el conjunto definido implícitamente por la ecuación $L(\mathbf{x}) = 0$. ¿Cómo se llama habitualmente este conjunto?

(a) $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$.

(b) $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$.

(c) $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

(d) $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 4 & 2 \\ 5 & -8 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

7. Dibuje las superficies definidas explícitamente por las funciones

(a) $f\begin{pmatrix} x \\ y \end{pmatrix} = 2 - x^2 - y^2$.

(b) $h\begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{1}{x^2 + y^2}$.

(c) $g(x, y) = \operatorname{sen} x$.

(d) $f(x, y) = 0$.

(e) $f(x, y) = e^{x+y}$.

(f) $g(x, y) = \begin{cases} 1 & \text{si } |x| < |y|, \\ 0 & \text{si } |x| \geq |y|. \end{cases}$

8. Dibuje las curvas definidas paramétricamente por las funciones

(a) $f(t) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} t + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad -\infty < t < \infty.$

(b) $f(t) = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}, \quad 0 \le t \le 1.$

(c) $f(t) = (2t, t), \quad -1 \le t \le 1.$

(d) $h(t) = (t, t, t^2), \quad -1 \le t \le 2.$

(e) $f(t) = \begin{pmatrix} 2t \\ |t| \end{pmatrix}, \quad -1 \le t \le 2.$

9. Dibuje las superficies definidas paramétricamente por las funciones

(a) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{cases} -\infty < u < \infty, \\ -\infty < v < \infty. \end{cases}$

(b) $g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos u \operatorname{sen} v \\ \operatorname{sen} u \operatorname{sen} v \\ \cos v \end{pmatrix}, \quad \begin{cases} 0 \le u \le 2\pi. \\ 0 \le v \le \pi/2. \end{cases}$

(c) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos u \cosh v \\ \operatorname{sen} u \cosh v \\ \operatorname{senh} v \end{pmatrix}, \quad \begin{cases} 0 \le u \le 2\pi, \\ -\infty < v < \infty. \end{cases}$

10. Dibuje los siguientes conjuntos de nivel definidos implícitamente

(a) $f(x, y) = x + y = 1.$

(b) $g(x, y) = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1.$

(c) $f(x, y) = (x^2 + y^2 + 1)^2 - 4x^2 = 0.$

(d) $f(x, y, z) = x + y + z = 1.$

(e) $xyz = 0.$

(f) $g(x, y, z) = x^2 - y^2 = 2.$

(g) $\begin{pmatrix} x - y \\ y + z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$

(h) $\begin{cases} 2x + y + z = 2, \\ x \qquad - z = 3. \end{cases}$

(i) $\begin{pmatrix} xyz \\ x + y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

(j) $xy + yz + zx = 1.$

11. Supongamos que la densidad por unidad de área de una película delgada está dada, en coordenadas rectangulares, por la fórmula $d(x, y) = x^2 + 2y^2 - x + 1$, para $-1 \le x \le 1$ y $-1 \le y \le 1$. Dibuje el conjunto de puntos en los cuales la película tiene densidad $\frac{7}{4}$.

12. Dibuje los campos vectoriales

(a) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix}$ para $-1 \le x \le 2, y = 0, y = 1.$

(b) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$ para $x^2 + y^2 \le 4.$

(c) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$ para $x^2 + y^2 \le 4.$

(d) $f\begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{1}{x^2 + y^2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ para $x^2 + y^2 \le 4.$

13. Sea f la transformación del plano xy en sí mismo definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ -x + y \end{pmatrix}.$$

Demuestre que f consiste en una expansión desde el origen en un factor de $\sqrt{2}$ y una rotación en un ángulo de $\pi/4$.

14. La función vectorial f está definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

¿Cuáles son las funciones coordenadas de f? Sea el plano xy su espacio dominio y el plano uv su espacio imagen.

(a) Encuentre la imagen del segmento de recta $y = x$ entre $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(b) Encuentre la imagen de la región definida por $0 < x$, $0 < y$ y $x^2 + y^2 < 1$.

(c) Encuentre el ángulo entre las imágenes de las rectas $y = 0$ e $y = (1/\sqrt{3})x$. [*Resp.* $\pi/3$.]

15. Una función vectorial f del plano xy en el plano uv, está definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ \dfrac{(x + y)^2}{4x} \end{pmatrix}, \quad x \ne 0.$$

¿Cuáles son las funciones coordenadas de f? Encuentre la imagen de la región acotada por las rectas $x = y$, $y = x - 8$, $x = -y$, $y = 8 - x$.

2. Funciones de una variable

Si un punto se mueve en un espacio vectorial, entonces su posición en el instante t puede ser descrita por una función vectorial f con valores $f(t)$. Por ejemplo, un punto que se mueve a lo largo de una recta en $\mathcal{R}^3$ estará en el punto

$$f(t) = t\mathbf{x}_1 + \mathbf{x}_0$$

en el instante t, donde $\mathbf{x}_1$ y $\mathbf{x}_0$ son puntos en $\mathcal{R}^3$. Más generalmente, una función f con valores en $\mathcal{R}^n$ puede ser dada en la forma

$$f(t) = (f_1(t), \ldots, f_n(t)),$$

donde las funciones coordenadas $f_1, \ldots, f_n$ describen las coordenadas de un punto de $\mathcal{R}^n$ en distintos instantes t.

Ejemplo 1. Si $\mathbf{x}_1 = (x_1, y_1, z_1)$ y $\mathbf{x}_0 = (x_0, y_0, z_0)$ son puntos de $\mathcal{R}^3$, entonces la función $\mathcal{R} \xrightarrow{f} \mathcal{R}^3$ definida por

$$\begin{aligned} f(t) &= t(x_1, y_1, z_1) + (x_0, y_0, z_0) \\ &= (tx_1 + x_0, ty_1 + y_0, tz_1 + z_0) \end{aligned}$$

dan una representación paramétrica para un punto en una recta. Análogamente, la función g tal que

$$g(t) = (t, t^2)$$

describe una curva en $\mathcal{R}^3$. Como las coordenadas $x = t$ e $y = t^2$ satisfacen la relación $y = x^2$, el punto (t, t^2) está siempre en la parábola de ecuación $y = x^2$, mostrada en la Fig. 12.

Podemos definir el límite de una función vectorial f con valores en $\mathcal{R}^n$ usando los límites de las funciones reales coordenadas f_k de f. Así, si

$$f(t) = (f_1(t), \ldots, f_n(t))$$

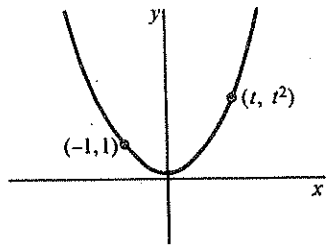


Figura 12

está definida en el intervalo $a < t < b$ que contiene a t_0, escribimos

$$\lim_{t \to t_0} f(t) = \left(\lim_{t \to t_0} f_1(t), \ldots, \lim_{t \to t_0} f_n(t) \right).$$

En forma análoga, una función con valores en $\mathcal{R}^n$ se dice continua si sus funciones coordenadas son continuas en su intervalo de definición.

Ejemplo 2. La función definida por $g(t) = (t, t^2)$ tiene un vector límite $(2, 4)$ en $t = 2$, ya que

$$\begin{aligned} \lim_{t \to 2} (t, t^2) &= \left(\lim_{t \to 2} t, \lim_{t \to 2} t^2 \right) \\ &= (2, 4). \end{aligned}$$

La función g es continua para todo t real, puesto que las funciones coordenadas t y t^2 lo son.

La idea intuitiva destrás de la noción de continuidad de una función vectorial es análoga a la que se tiene para las funciones reales: los valores de la función no deben cambiar abruptamente. Por ahora solo consideraremos funciones continuas $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, donde $g(t)$ está definida en un intervalo $a < t < b$. Primero definiremos la derivada de g y luego veremos que esto conduce a una definición de la recta tangente a la curva γ definida como la imagen g.

La función g tiene una **derivada** $g'(t)$ en el punto t del intervalo (a, b) si

$$g'(t) = \lim_{h \to 0} \frac{g(t + h) - g(t)}{h}.$$

Si este límite existe para todo t en (a, b), entonces $g'(t)$ determina una nueva función $\mathcal{R} \xrightarrow{g'} \mathcal{R}^n$, igual que en el caso $n = 1$. A menudo, la derivada también se denota por dg/dt.

Ejemplo 3. Sea $g(t) = (t^2, t^3)$. Entonces

$$\begin{aligned} \lim_{h \to 0} \frac{g(t + h) - g(t)}{h} &= \lim_{h \to 0} \frac{1}{h} \begin{pmatrix} (t + h)^2 - t^2 \\ (t + h)^3 - t^3 \end{pmatrix} \\ &= \lim_{h \to 0} \begin{pmatrix} \dfrac{(t + h)^2 - t^2}{h} \\ \dfrac{(t + h)^3 - t^3}{h} \end{pmatrix}. \end{aligned}$$

Como $\lim_{h \to 0} ((t + h)^2 - t^2)/h = 2t$ y $\lim_{h \to 0} ((t + h)^3 - t^3)/h = 3t^2$, el límite $g'(t)$ existe y $g'(t) = (2t, 3t^2)$.

Este ejemplo sugiere que una función $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ tiene derivada en el punto

t si, y solo si, cada función coordenada de g tiene derivada en ese punto. Esto es verdad y, en efecto, tenemos

2.1 Si $g(t) = \begin{pmatrix} g_1(t) \\ \cdot \\ \cdot \\ \cdot \\ g_n(t) \end{pmatrix}$, entonces $g'(t) = \begin{pmatrix} g_1'(t) \\ \cdot \\ \cdot \\ \cdot \\ g_n'(t) \end{pmatrix}$, donde las derivadas

$g_k'(t)$ son derivadas ordinarias de funciones reales de una variable real. Este resultado es una consecuencia inmediata de la relación entre el límite de una función vectorial y los límites de sus funciones coordenadas.

Ejemplo 4. Si

$$g(t) = \begin{pmatrix} \cos t \\ \operatorname{sen} t \end{pmatrix}, \text{ entonces } g'(t) = \begin{pmatrix} -\operatorname{sen} t \\ \cos t \end{pmatrix}.$$

Si

$$h(t) = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}, \text{ entonces } h'(t) = \begin{pmatrix} 1 \\ 2t \\ 3t^2 \end{pmatrix}.$$

La Fig. 13 muestra claramente que el vector $g(t + h) - g(t)$ tiene una dirección que, cuando h tiende a 0, debería tender a algo que nos gustaría llamar la dirección tangente a la curva γ en $g(t)$. Sin embargo, como g es continua

$$\lim_{h \to 0} g(t + h) - g(t) = 0,$$

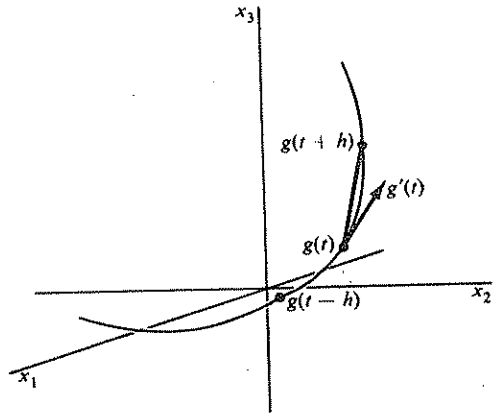


Figura 13

y el vector cero que obtenemos como límite, no tiene dirección. Esta dificultad se resuelve en la mayor parte de los ejemplos dividiendo por h antes de hacer tender a cero. Nótese que la división por h no cambia la dirección de $g(t + h) - g(t)$ si h es positivo y la invierte si es negativo. Una mirada a la Fig. 13 muestra que esto es deseable para nuestros propósitos. (¿Qué fallaría si dividiéramos por $|h|$?) Si la derivada $g'(t)$ existe y no es cero, se llama el **vector tangente** a γ en $g(t)$. Por supuesto, todo múltiplo distinto de cero de $g'(t)$ es llamado *un* **vector tangente** y la recta que determina $g'(t)$ y que pasa por $g(t)$ se llama la **recta tangente** a γ en $g(t)$. Por tanto, si $g(t_0)$ es un punto de una curva, la recta tangente a $g(t_0)$ tendrá una representación paramétrica de la forma

$$tg'(t_0) + g(t_0).$$

El vector tangente $g'(t_0)$ se representa generalmente con su cola en $g(t_0)$, como en la Fig. 13.

Ejemplo 5. La circunferencia definida paramétricamente por $g(t) = (\cos t, \operatorname{sen} t)$ tiene el vector tangente en $g(t_0)$ dado por $g'(t_0) = (-\operatorname{sen} t_0, \cos t_0)$. En particular, el vector tangente en $g(\pi/4) = (1/\sqrt{2}, 1/\sqrt{2})$ es $g'(\pi/4) = (-1/\sqrt{2}, 1/\sqrt{2})$. Por tanto, la recta tangente a la circunferencia en $(1/\sqrt{2}, 1/\sqrt{2})$ tiene la representación paramétrica $(x, y) = t(-1/\sqrt{2}, 1/\sqrt{2}) + (1/\sqrt{2}, 1/\sqrt{2})$. Esta recta está indicada en la Fig. 14, donde también se muestran otros vectores tangentes, cada uno de los cuales tiene largo $|g'(t_0)| = 1$.

Para la curva espiral definida por $f(t) = (\cos t, \operatorname{sen} t, t)$, tenemos que $f'(0) = (0, 1, 1)$. La recta tangente a la espiral en $(1, 0, 0)$ puede ser representada por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

y, en este caso, el vector tangente $f'(0)$ tiene largo $\sqrt{2}$.

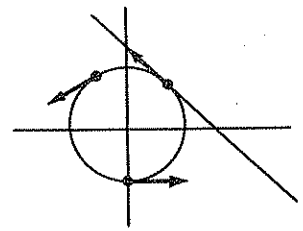


Figura 14

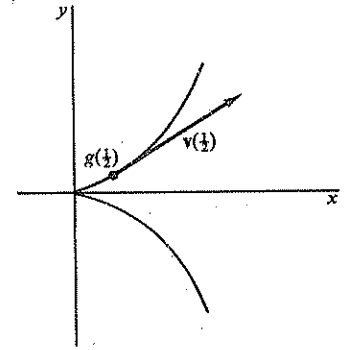


Figura 15

Una razón para destacar $g'(t)$ y no un múltiplo de él y llamarlo *el* vector tangente es que, a menudo queremos considerar el parámetro t como el tiempo y entonces $g(t)$ representa el camino que sigue un punto que se mueve en $\mathfrak{R}^n$. En esta interpretación, el largo euclidiano $|g'(t)|$ es la definición natural de la rapidez del movimiento a lo largo del camino γ descrito por $g(t)$. Para justificar el uso del término «rapidez», observemos que, si h es pequeño, entonces el número $|g(t+h) - g(t)|/|h|$ está cerca de la rapidez promedio en el intervalo de t a $t+h$. Además, es fácil probar que si $g'(t)$ existe, entonces

$$\lim_{h\to 0} \frac{|g(t+h) - g(t)|}{|h|} = |g'(t)| .$$

En efecto, por la desigualdad del Ejercicio 12 de la Sec. 5 del Cap. 1,

$$\left| \frac{|g(t+h) - g(t)|}{|h|} - |g'(t)| \right| \leq \left| \frac{g(t+h) - g(t)}{h} - g'(t) \right|,$$

que tiende a cero cuando h tiende a cero. Por tanto, $|g'(t)|$ es el límite de la rapidez promedio en intervalos de tiempos arbitrariamente pequeños. Por otra razón, la función real v, definida por $v(t) = |g'(t)|$, se llama la **velocidad** de g y el vector $v(t) = g'(t)$ se llama el **vector velocidad** en el punto $g(t)$. El vector $\mathbf{v}(t)$ es, evidentemente, el vector tangente a γ en $g(t)$, siempre que $\mathbf{v}(t) \neq 0$.

Ejemplo 6. Si un punto se mueve en el plano, de modo que, en el instante t, su posición es $g(t) = (t^2, t^3)$, entonces el vector velocidad es $\mathbf{v}(t) = (2t, 3t^2)$ y $v(t) = \sqrt{4t^2 + 9t^4}$. En particular, $\mathbf{v}(0) = 0$. El camino seguido por g se muestra en la Fig. 15 para $-1 \leq t \leq 1$, donde es conveniente notar que las coordenadas de un punto en ese camino satisfacen la ecuación $y^2 = x^3$.

El hecho que el vector tangente tiende a 0 cuando $g(t)$ se acerca al origen es un reflejo del hecho de que si la velocidad cambia continuamente, entonces la rapidez debe ser cero en un punto en que hay un cambio abrupto de la dirección de movimiento. Es así cómo la parametrización describe también la situación física. Sin embargo, esta parametrización no sirve para definir una recta tangente al camino en el origen.

Ejemplo 7. Sea $f(t) = (\cos t, \operatorname{sen} t, t)$, como en la segunda parte del Ejemplo 5. Entonces $\mathbf{v}(t) = (-\operatorname{sen} t, \cos t, 1)$. Por tanto, el vector velocidad es siempre perpendicular al vector $(\cos t, \operatorname{sen} t, 0)$. La rapidez en cualquier instante t es $v(t) = |\mathbf{v}(t)| = \sqrt{2}$.

Damos a continuación algunas fórmulas útiles que son válidas cuando f y g tienen derivada vectorial en un intervalo (a, b) y φ es una función real diferenciable definida en el mismo intervalo.

2.2 $(f+g)' = f' + g'$, $(cf)' = cf'$, c **constante**

2.3 $(\varphi f)' = \varphi f' + \varphi' f$, φ real

2.4 $(f \cdot g)' = f \cdot g' + f' \cdot g.$

2.5 $(f(u))' = u' f'(u),$

donde u es una función real diferenciable de una variable, cuya imagen está en (a, b).

Estas fórmulas pueden ser demostradas escribiendo f y g en términos de sus coordenadas, luego aplicando las fórmulas correspondientes para funciones reales y, finalmente, aplicando la Ec. 2.1. Por ejemplo, la demostración de 2.5, que es una versión de la regla de la cadena para la derivación, es como sigue:

$$\begin{aligned}(f(u))' &= (f_1(u), \ldots, f_n(u))' \\ &= ([f_1(u)]', \ldots, [f_n(u)]') \\ &= (f_1'(u)u', \ldots, f_n'(u)u') = u'f'(u).\end{aligned}$$

Si $\mathfrak{R} \xrightarrow{g} \mathfrak{R}^n$ tiene derivada $\mathfrak{R} \xrightarrow{g'} \mathfrak{R}^n$, entonces podemos considerar la derivada de g' que denotamos por g''. Tenemos así $\mathfrak{R} \xrightarrow{g''} \mathfrak{R}^n$, a pesar de que g'' puede estar definida en menos puntos que g'. Escribimos también d^2g/dt^2 en lugar de g''. Podemos continuar así considerando derivadas de orden superior.

Ejemplo 8. Sea $\mathfrak{R} \xrightarrow{g} \mathfrak{R}^3$ la descripción de un camino en $\mathfrak{R}^3$ con vector velocidad $\mathbf{v}(t)$ y rapidez $v(t)$ en cada punto $g(t)$. Entonces $\mathbf{t}(t) = (1/v(t))\mathbf{v}(t)$ es un vector tangente de largo 1, siempre que $v(t) \neq 0$. Si suponemos que g' tiene una derivada, definimos el **vector aceleración** en $g(t)$ como $\mathbf{a}(t) = g''(t)$. El significado físico de $\mathbf{a}(t)$ es que si $g(t)$ describe el movimiento de una partícula de masa constante m, entonces $m\mathbf{a}(t)$ es el **vector fuerza** $F(t)$ que actúa sobre la partícula. Si denotamos por $a(t)$ el largo del vector $\mathbf{a}(t)$, entonces $a(t)$ es la **magnitud de la aceleración** y $ma(t)$ es la **magnitud de la fuerza** que actúa sobre la partícula.

Si $\mathbf{t}(t)$ es un vector tangente unitario en $g(t)$, la ecuación $g'(t) = v(t)\mathbf{t}(t)$ implica que $\mathbf{a} = (v\mathbf{t})'$. Aplicando la Ec. 2.3, tenemos que $\mathbf{a} = v'\mathbf{t} + v\mathbf{t}'$. Por tanto, si $\mathbf{t}'(t) = 0$, entonces el vector aceleración y, por tanto, también el vector fuerza, en $g(t)$ tiene la misma dirección que la dirección de movimiento o la dirección opuesta. Por otra parte, si $\mathbf{t}'(t) \neq 0$, podemos definir el vector unitario $\mathbf{n}$ por $\mathbf{n}(t) = \mathbf{t}'(t)/|\mathbf{t}'(t)|$ y, por tanto, el vector aceleración puede ser escrito en la forma

$$\mathbf{a} = v'\mathbf{t} + v\,|\mathbf{t}'|\,\mathbf{n}.$$

Esta ecuación expresa la aceleración $\mathbf{a}(t)$ en cada punto $g(t)$ en términos del par ortonormal de vectores $\mathbf{t}(t)$ y $\mathbf{n}(t)$. En efecto, $|\mathbf{t}| = |\mathbf{n}| = 1$ por la definición de estos vectores y aplicando la Ec. 2.4 a la ecuación $\mathbf{t} \cdot \mathbf{t} = 1$ se tiene que

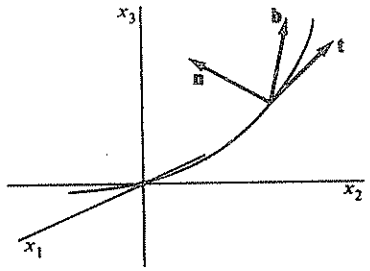


Figura 16

$\mathbf{t}' \cdot \mathbf{t} = 0$. Por tanto, por definición de $\mathbf{n}$, $\mathbf{t} \cdot \mathbf{n} = 0$.

El par $\mathbf{t}(t)$, $\mathbf{n}(t)$, se representa en el punto $g(t)$ en la Fig. 16. El tercer vector unitario $\mathbf{b}(t)$ mostrado allí está definido por $\mathbf{b} = \mathbf{t} \times \mathbf{n}$ y se llama el vector **binormal** al camino, mientras que $\mathbf{n}$ recibe el nombre de **normal principal.** Es así cómo cualquier vector, asociado naturalmente a cada punto $g(t)$ del camino, puede ser expresado como una combinación lineal del triple $\{\mathbf{t}(t), \mathbf{n}(t), \mathbf{b}(t)\}$ que cambia al ir de un punto a otro a lo largo del camino.

EJERCICIOS

1. Sea $g(t) = (e^t, t)$, dibuje en $\mathfrak{R}^2$ la curva descrita por g y los vectores tangentes $g'(0)$ y $g'(1)$.
2. Sea $f(t) = (t, t^2, t^3)$ para $0 \leq t \leq 1$.
   (a) Dibuje la curva descrita por f en $\mathfrak{R}^3$ y la recta tangente en $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8})$.
   (b) Calcule $|f'(t)|$.
   (c) Si $f(t) = (t, t^2, t^3)$ para todo t real, encuentre todos los puntos de la curva descrita por f en que el vector tangente es paralelo al vector $(4, 4, 3)$. ¿Hay puntos en los cuales la tangente es perpendicular a $(4, 4, 3)$?
3. Dibuje la curva representada por $(x, y) = (t^3, t^5)$ y demuestre que esta parametrización no define un vector tangente en el origen. Encuentre una parametrización de la curva que define un vector tangente en el origen.
4. Demuestre que la curva descrita por $g(t) = (\operatorname{sen} 2t, 2 \operatorname{sen}^2 t, 2 \cos t)$ yace en una esfera con centro en el origen de $\mathfrak{R}^3$. Encuentre el largo del vector velocidad $\mathbf{v}(t)$ y demuestre que la proyección de este vector en el plano xy tiene largo constante.
5. (a) Demuestre que si $\mathfrak{R} \xrightarrow{\mathbf{v}} \mathfrak{R}^3$ es continua para $t \geq 0$, entonces existe un único camino g que pasa por un punto dado $g(0)$ de $\mathfrak{R}^3$ y que tiene vector velocidad $\mathbf{v}(t)$ en $g(t)$.
   (b) Demuestre que una función continua $\mathfrak{R} \xrightarrow{\mathbf{a}} \mathfrak{R}^3$ definida para $t \geq 0$ determina un camino único en $\mathfrak{R}^3$ que tiene vector aceleración $\mathbf{a}(t)$, siempre que se especifique su punto inicial $g(0)$ y su velocidad inicial $\mathbf{v}(0)$.
6. Supongamos que un blanco se mueve con rapidez constante $v > 0$ en un camino circular de radio r y que un cohete lo persigue con velocidad también constante v, partiendo desde el centro de la circunferencia, manteniéndose siempre en la recta que une al centro y al blanco. ¿Cuándo da en el blanco el cohete?

7. Demuestre (a) 2.2, (b) 2.3, (c) 2.4, de la Sec. 2.
8. Demuestre que si f es una función vectorial diferenciable y distinta de cero para $a < t < b$, entonces
   (a) $f \cdot \dfrac{df}{dt} = |f| \dfrac{d|f|}{dt}$.
   (b) $|f|$ es constante si, y solo si, $f \cdot f' = 0$.
9. Consideremos la ecuación diferencial
$$\mathbf{x}'' + a\mathbf{x}' + b\mathbf{x} = 0$$
que debe tener como solución una función $\mathbf{x}(t)$ con valores en $\mathfrak{R}^n$ definida en algún intervalo. Supongamos que a y b son constantes. Demuestre que si la ecuación real $r^2 + ar + b = 0$ tiene soluciones distintas r_1 y r_2, entonces la ecuación diferencial tiene una solución de la forma
$$\mathbf{x}(t) = \mathbf{c}_1 e^{r_1 t} + \mathbf{c}_2 e^{r_2 t},$$
donde $\mathbf{c}_1$ y $\mathbf{c}_2$ son dos vectores constantes en $\mathfrak{R}^n$. ¿Qué sucede si $r_1 = r_2$?
10. Demuestre que $(f \times g)' = (f \times g') + (f' \times g)$, donde f y g toman valores en $\mathfrak{R}^3$ y son diferenciables en algún intervalo.
11. Demuestre que si $\mathfrak{R} \xrightarrow{g} \mathfrak{R}^n$ tiene derivada y $g'(t) = 0$ para $a < t < b$, entonces $g(t)$ es un vector constante en todo el intervalo. [*Indicación*: Aplique el teorema del valor medio a cada función coordenada.]
12. Suponga que la función diferenciable $g(t)$ representa la posición en $\mathfrak{R}^3$, en el instante t, de una partícula de masa posiblemente variable $m(t)$. La función vectorial $P(t) = m(t)\mathbf{v}(t)$ se llama el **momento lineal** de la partícula. El **vector fuerza** es $F(t) = (m(t)\mathbf{v}(t))'$. El **momento angular** con respecto al origen es $L(t) = g(t) \times P(t)$ y el **torque** con respecto del origen es $N(t) = g(t) \times F(t)$.
   (a) Demuestre que si F es idénticamente cero, entonces P es constante. Esta afirmación se conoce con el nombre de ley de conservación del momento lineal.
   (b) Demuestre que $L'(t) = N(t)$ y que, por tanto, si N es idénticamente cero, entonces L es constante. Esto se conoce con el nombre de ley de conservación del momento angular.
13. Demuestre que si una partícula tiene vector aceleración $\mathbf{a}(t)$ en el instante t y $v(t) \neq 0$, entonces $v' = \mathbf{t} \cdot \mathbf{a}$, donde $\mathbf{t}$ es el vector unitario $(1/v)\mathbf{v}$.
14. Sea $\mathfrak{R} \xrightarrow{f} \mathfrak{R}$ una función definida para $a \leq t \leq b$. Si las funciones coordenadas $f_1, \ldots, f_n$ de f son integrables, podemos definir la integral de f en el intervalo $[a, b]$ por
$$\int_a^b f(t)\,dt = \left(\int_a^b f_1(t)\,dt, \ldots, \int_a^b f_n(t)\,dt\right).$$
   (a) Si $f(t) = (\cos t, \operatorname{sen} t)$ para $0 \leq t \leq \pi/2$, calcule $\displaystyle\int_0^{\pi/2} f(t)\,dt$.
   (b) Si $g(t) = (t, t^2, t^3)$ para $0 \leq t \leq 1$, calcule $\displaystyle\int_0^1 g(t)\,dt$.

15. Si $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$ y $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ son integrables en $[a, b]$, demuestre usando las propiedades correspondientes para funciones reales, que

$$\int_a^b kf(t)\,dt = k\int_a^b f(t)\,dt, \qquad k \text{ número real}$$

$$\int_a^b (f(t) + g(t))\,dt = \int_a^b f(t)\,dt + \int_a^b g(t)\,dt,$$

donde las integrales han sido definidas en el ejercicio anterior.

16. Si $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$ está definida para $a \le t \le b$ y f' es continua en este mismo intervalo, demuestre la siguiente extensión del teorema fundamental del cálculo:

$$\int_a^b f'(t)\,dt = f(b) - f(a).$$

17. Si $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$ es continua en $[a, b]$,

(a) Demuestre que

$$\int_a^b \mathbf{k} \cdot f(t)\,dt = \mathbf{k} \cdot \int_a^b f(t)\,dt,$$

donde $\mathbf{k}$ es un vector constante.

(b) Demuestre que

$$\left|\int_a^b f(t)\,dt\right| \le \int_a^b |f(t)|\,dt.$$

[*Indicación*: Por la desigualdad de Cauchy-Schwarz

$$f(u) \cdot \int_a^b f(t)\,dt \le |f(u)|\left|\int_a^b f(t)\,dt\right|, \quad \text{para todo } u.$$

Integre con respecto a u y aplique el resultado de la parte (a).]

3. Longitud de arco

La definición del largo de vectores puede ser usada para definir el largo de una curva parametrizada γ. Suponemos que γ está descrita por una función continua $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, donde el dominio de g es el intervalo cerrado $a \le t \le b$. Por tanto, γ es la imagen por g de $[a, b]$. Para cada conjunto finito P de números $a = t_0 < t_1 < \cdots < t_K = b$, hay puntos $g(t_k)$, $k = 0, \ldots, K$, en γ. Unimos estos puntos, en ese orden, por un camino poligonal como lo muestra la Fig. 17. El largo del segmento k de esta aproximación poligonal de γ es $|g(t_k) - g(t_{k-1})|$ y el largo total de este polígono es

$$l(P) = \sum_{k=1}^{K} |g(t_k) - g(t_{k-1})|.$$

Denotemos por $l(\gamma)$ al supremo de los números $l(P)$. Por supuesto, esto será infinito si el conjunto de todos los $l(P)$ no es acotado. Si $l(\gamma)$ es finito, entonces se dice que γ es **rectificable** y $l(\gamma)$ se llama su **largo.** Es claro que la definición de $l(\gamma)$ depende de la función g que describe a γ. Esto es razonable, ya que queremos tomar en consideración el hecho de que alguna parte de γ puede ser recorrida más de una vez por g. En la práctica, es muy corriente que esto sea deseable. Si g no es inyectiva, podremos escribir $l(g)$ en lugar de $l(\gamma)$ cuando queramos poner énfasis en la dependencia de g.

Generalmente, es muy complicado calcular la longitud de un camino a partir de su definición. Sin embargo, si γ es parametrizada por una función g tal que el vector tangente $g'(t)$ varía continuamente con t, entonces $l(\gamma)$ es finita y es igual a $\int_a^b |g'(t)|\,dt$. Es más, es suficiente suponer que g es **lisa por pedazos,** lo que significa que g' es continua en cada uno de un número finito de intervalos $a \le t \le b$. Esto nos permite hacer el cálculo de algunas curvas, para las cuales el vector tangente tiene un cambio brusco de dirección, como lo muestra la Fig. 17.

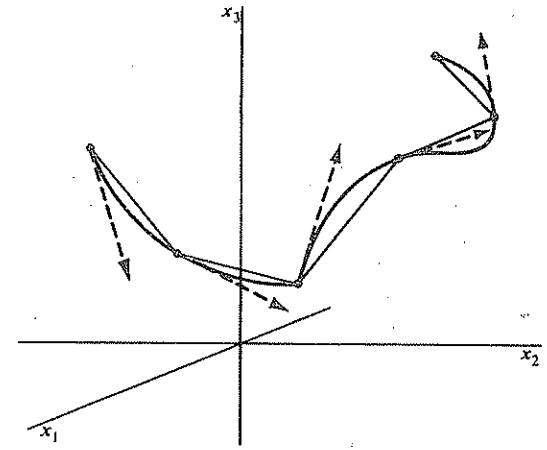


Figura 17

3.1 Teorema

Sea γ una curva parametrizada por una función lisa por pedazos $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, definida para $a \le t \le b$. Entonces $l(\gamma)$ es finita y

$$l(\gamma) = \int_a^b |g'(t)|\,dt.$$

La demostración de este teorema se dará en la Sec. 3 del Apéndice. Aquí solo daremos un argumento que hace más plausible esta fórmula. Ya que, por definición, la existencia de $g'(t_{k-1})$ significa que

$$\lim_{t_k \to t_{k-1}} \frac{g(t_k) - g(t_{k-1})}{t_k - t_{k-1}} = g'(t_{k-1}),$$

tenemos que

$$g(t_k) - g(t_{k-1}) - (t_k - t_{k-1})g'(t_{k-1}) = (t_k - t_{k-1})Z(t_k - t_{k-1}),$$

donde $Z(t_k - t_{k-1})$ satisface

$$\lim_{t_k \to t_{k-1}} Z(t_k - t_{k-1}) = 0.$$

Por tanto, $(t_k - t_{k-1})g'(t_{k-1})$ se hace progresivamente una mejor aproximación de $g(t_k) - g(t_{k-1})$ a medida que $t_k - t_{k-1}$ se hace pequeño. Es así como podemos aproximar

$$l(P) = \sum_{k=1}^{K} |g(t_k) - g(t_{k-1})|$$

por

$$\sum_{k=1}^{K} |g'(t_{k-1})|\,(t_k - t_{k-1}).$$

(Algunos de los vectores tangentes $(t_k - t_{k-1})g'(t_{k-1})$ se muestran en la Fig. 17.) Ahora bien, si g' es continua, también lo es $|g'|$ y haciendo $m(P) = \max_{1 \le k \le K} (t_k - t_{k-1})$ tenemos que

$$\lim_{m(P) \to 0} \sum_{k=1}^{K} |g'(t_{k-1})|\,(t_k - t_{k-1}) = \int_a^b |g'(t)|\,dt.$$

lo cual termina el argumento,

Una curva en $\mathcal{R}^n$ se describe generalmente mediante funciones coordenadas. Así, si γ está definida, para $a \le t \le b$ por $g(t) = (g_1(t), g_2(t), g_3(t))$, entonces $|g'(t)| = \sqrt{(g_1(t))^2 + (g_2(t))^2 + (g_3(t))^2}$ y, por tanto,

$$l(g) = \int_a^b \sqrt{(g_1'(t))^2 + (g_2'(t))^2 + (g_3'(t))^2}\,dt,$$

En el caso de $\mathcal{R}^n$ se tiene una fórmula similar. Por ejemplo, la curva espiral en $\mathcal{R}^3$ definida por $g(t) = (\cos t,\ \operatorname{sen} t,\ t)$, para $0 \le t \le 1$, tiene largo

$$\begin{aligned} l(g) &= \int_0^1 \sqrt{(-\operatorname{sen} t)^2 + (\cos t)^2 + 1}\,dt \\ &= \sqrt{2}. \end{aligned}$$

Ejemplo 1. La curva plana definida por $g(t) = (t^2, |t|)$ para $-1 \le t \le 1$ se muestra en la Fig. 18. Como es lisa por pedazos, $l(g)$ es finito. Tenemos que

$$g'(t) = \begin{cases} (2t, -1), & -1 \le t < 0 \\ (2t, 1), & 0 < t \le 1, \end{cases}$$

de modo que $|g'(t)| = \sqrt{4t^2 + 1}$ para $-1 \le t \le 1$, $t \ne 0$. Por tanto,

$$\begin{aligned} l(g) &= \int_{-1}^{1} \sqrt{4t^2 + 1}\,dt \\ &= \sqrt{5} + (\tfrac{1}{2}) \log (2 + \sqrt{5}). \end{aligned}$$

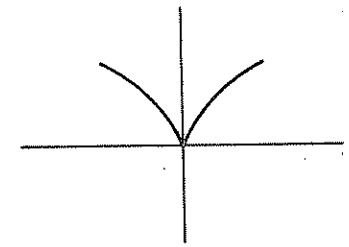


Figura 18

Ejemplo 2. El gráfico γ de una función real continua f, definida para $a \le x \le b$, es una curva en $\mathcal{R}^2$ que puede ser descrita paramétricamente por la función inyectiva

$$g(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix}, \qquad a \le x \le b.$$

Si f' es continua, entonces también lo es $|g'(x)| = \sqrt{1 + (f'(x))^2}$ y la fórmula para encontrar $l(\gamma)$ es, en este caso,

$$l(\gamma) = \int_a^b \sqrt{1 + (f'(x))^2}\,dx.$$

Si γ es una curva lisa por pedazos, descrita por una función inyectiva $g(t)$, $a \le t \le b$, entonces podemos considerar que $l(\gamma)$ representa la masa total de γ, si se supone que γ tiene una densidad igual a 1 en cada punto. En forma más general, si p es una función real definida en γ, podemos formar la integral

$$\int_a^b p(g(t))\,|g'(t)|\,dt,$$

que, si existe, llamamos la integral de p en γ. En particular, si p es una función no negativa, que puede ser considerada como la densidad por unidad de longitud de una distribución de masa en γ. En este caso, la integral es la definición de la **masa total** m de la distribución.

Ejemplo 3. Consideremos una vuelta completa de la curva espiral definida por $g(t) = (\cos t,\ \operatorname{sen} t,\ t)$ para $0 \le t \le 2\pi$. Supongamos que la densidad de la curva en un punto es igual al cuadrado de la distancia de este punto al punto medio $\mathbf{q}$ del eje de la espiral. En nuestra descripción de la curva, el punto $\mathbf{q}$ es $(0, 0, \pi)$. Por tanto, la densidad en $g(t)$ es igual a

$$\begin{aligned} |g(t) - \mathbf{q}|^2 &= \cos^2 t + \operatorname{sen}^2 t + (t - \pi)^2 \\ &= 1 + (t - \pi)^2. \end{aligned}$$

Hemos visto que $|g'(t)| = \sqrt{2}$. Por tanto, la masa total de esta distribución está dada por

$$m = \int_0^{2\pi} [1 + (t - \pi)^2]\sqrt{2}\, dt$$

$$= \sqrt{2}\int_{-\pi}^{\pi} (1 + t^2)\, dt = (\tfrac{2}{3})\sqrt{2}\,(3\pi + \pi^3).$$

Si $g(t)$, con $a \le t \le b$, define una curva lisa γ con vector tangente distinto de cero de cada punto, entonces la función $s(t)$, definida por

3.2 $$s(t) = \int_a^t |g'(u)|\, du$$

es el largo de la parte de γ que corresponde al intervalo de a a t. Como $|g'(t)|$ es positivo, la función $s(t)$ es estrictamente creciente. Es así como $s(t)$ es una función biyectiva del intervalo $a \le t \le b$ con el intervalo $0 \le s \le l(\gamma)$ y, por tanto, tiene una función inversa. Denotamos esta función inversa por $r(s)$. Ahora formamos la función vectorial $h(s) = g(r(s))$, que describe la curva γ descrita por g, pero con una nueva parametrización, en la cual la variable es s, con $0 \le s \le l(\gamma)$, que representa el largo del camino γ desde $h(0)$ a $h(s)$. La curva γ se dice entonces parametrizable por el largo del arco.

Ejemplo 4. Sea $g(t) = (t, t^2)$, para $0 \le t \le 1$. Entonces

$$s(t) = \int_0^t \sqrt{1 + 4u^2}\, du$$

$$= \tfrac{1}{2}t\sqrt{1 + 4t^2} + \tfrac{1}{4}\log\,(2t + \sqrt{1 + 4t^2})$$

Como $s'(t) = \sqrt{1 + 4t^2} > 0$, $s(t)$ es estrictamente creciente, tiene un inverso.

El Ejemplo 4 muestra que $r(s)$, el inverso de $s(t)$, puede ser muy complicado para calcular. Sin embargo, su uso tiene varias ventajas teóricas. Por ejemplo, si la curva γ está parametrizada por el largo de arco, entonces la integral de una función real p a lo largo de γ toma una forma más simple. Si γ está definida por la función $g(t)$ con $a \le t \le b$, entonces la nueva parametrización está dada por la función $h(s)$, definida para $0 \le s \le l(\gamma)$, donde $h(s) = g(r(s))$. Si hacemos el cambio de variable en la integral de p a lo largo de γ, obtenemos, ya que $s'(t) = |g'(t)|$,

$$\int_a^b p(g(t))\,|g'(t)|\, dt = \int_a^b p(g(t))s'(t)\, dt$$

$$= \int_0^{l(\gamma)} p(h(s))\, ds.$$

La expresión $|g'(t)\, dt|$, o su contrapartida más sencilla ds, se llama a menudo el **elemento de largo de arco** de la curva γ. De la Ec. 3.2, que muestra la relación entre s y t, podemos deducir una relación entre ambas

3.3 $$\frac{ds}{dt}(t) = |g'(t)| = v(t).$$

Para demostrar esto, se derivan con respecto de t ambos términos de 3.2 y se usa la *definición* de la rapidez $v(t)$ como $|g'(t)|$. La Ec. 3.3 da una nueva justificación para la definición de rapidez: la derivada del largo de arco con respecto al tiempo es la velocidad.

Ejemplo 5. Recordamos que la ecuación

$$\mathbf{a}(t) = \mathbf{v}'(t)\mathbf{t}(t) + v(t)\,|\mathbf{t}'(t)|\,\mathbf{n}(t), \tag{1}$$

deducida en el Ejemplo 8 de la sección anterior, expresa el vector aceleración de una curva en cada punto de ella como una combinación lineal de dos vectores unitarios, uno tangente y el otro perpendicular a la curva. El coeficiente de $\mathbf{n}(t)$ puede ser escrito en una forma más interesante. Si escribimos $r(s)$ para el inverso de $s(t)$, por la Ec. 2.5, tenemos que

$$\frac{d}{ds}\mathbf{t}(r(s)) = \mathbf{t}'(r(s))\frac{dr}{ds}(s). \tag{2}$$

En general, para funciones inversas r y s, tenemos que $r(s(t)) = t$ y, por la regla de la cadena,

$$\frac{dr}{ds}(s(t))\frac{ds}{dt}(t) = 1.$$

De la Ec. 3.3, se tiene que

$$\frac{dr}{ds}(s) = \frac{1}{v(t)}. \tag{3}$$

Entonces, de las Ecs. (2) y (3), se obtiene

$$\mathbf{t}'(t) = v(t)\frac{d}{ds}\mathbf{t}(t)$$

Como $v(t) \ge 0$, si tomamos el largo de ambos términos, tenemos que

$$|\mathbf{t}'(t)| = v(t)\left|\frac{d}{ds}\mathbf{t}(t)\right|.$$

El factor

$$\kappa(t) = \left|\frac{d}{ds}\mathbf{t}(t)\right| = \frac{|\mathbf{t}'(t)|}{v(t)}$$

se llama la **curvatura** de la curva en el punto que corresponde al tiempo t. La Ec. (1) toma ahora la forma

3.4 $$\mathbf{a}(t) = v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\mathbf{n}(t).$$

Los términos de esta suma se llaman las **componentes tangencial** y **normal** (o **centrípeta**) de la aceleración respectiva y los factores numéricos se denotan a veces por $a_t = v'$ y $a_n = v^2\kappa$.

De la Ec. 3.4, se pueden deducir inmediatamente varias cosas acerca del vector aceleración **a**. Si la rapidez v es una constante v_0, entonces $v' = 0$ y $\mathbf{a}(t) = v_0^2\kappa(t)\mathbf{n}(t)$. Esto significa que **a** es perpendicular a la curva y que su largo

$v_0^2\kappa(t)$ depende solo de la curvatura κ. En el otro extremo, si el movimiento es a lo largo de una línea recta, entonces el vector tangente $\mathbf{t}$ es constante, de modo que

$$\frac{d\mathbf{t}}{ds}(r(s)) = 0\,.$$

Por tanto, $\kappa = 0$ y tenemos que $\mathbf{a}(t) = v'(t)\mathbf{t}(t)$, lo cual muestra que el vector aceleración tiene la misma dirección que la tangente o la dirección opuesta, de acuerdo con el signo de v'. El Ejercicio 10 muestra que κ, la curvatura del camino, mide la rapidez con que varía el vector tangente $\mathbf{t}$.

EJERCICIOS

1. Encuentre el largo de las siguientes curvas:

(a) $(x, y) = (t, \log\cos t)$, $0 \le t \le 1$. [*Resp*: $\log(\sec 1 + \tan 1)$.]
(b) $(x, y) = (t^2, \frac{2}{3}t^3 - \frac{1}{2}t)$, $0 \le t \le 2$. [*Resp*: $\frac{19}{3}$.]
(c) $y = x^{3/2}$, $0 \le x \le 5$.
(d) $g(t) = (6t^2, 4\sqrt{2}\,t^3, 3t^4)$, $-1 \le t \le 2$. [*Resp*: 81.]

2. (a) Escriba la integral que expresa el largo de arco de la elipse

$$(x, y) = (a\cos t, b\,\mathrm{sen}\,t),\ 0 \le t \le 2\pi.$$

(b) Demuestre que el cálculo de la integral de la parte (a) se puede reducir al cálculo de la integral elíptica de la forma

$$\int_0^{\pi/2} \sqrt{1 - k^2\mathrm{sen}^2\,\theta}\; d\theta.$$

(c) Usando una tabla de integrales elípticas o usando un cálculo numérico, encuentre un valor aproximado del largo de arco de la elipse con $a = 1$ y $b = 2$.

3. Supongamos que la curva γ está definida por las dos funciones continuamente diferenciables

$$f(t), \qquad a \le t \le b,$$
$$g(u), \qquad \alpha \le u \le \beta.$$

Estas funciones se llaman parametrizaciones **equivalentes** de γ si existe una función φ continuamente diferenciales tales que

$$a = \varphi(\alpha) \quad \text{y} \quad b = \varphi(\beta),$$
$$f(\varphi(u)) = g(u), \qquad \alpha \le u \le \beta,$$
$$\varphi'(u) > 0, \qquad \alpha < u < \beta.$$

(a) Demuestre que dos parametrizaciones equivalentes de γ asignan el mismo largo a γ.
(b) Demuestre que

$$(x, y) = (\cos t, \mathrm{sen}\,t), \qquad 0 \le t \le \frac{\pi}{2}$$

y

$$(x, y) = \left(\frac{1 - u^2}{1 + u^2}, \frac{2u}{1 + u^2}\right), \qquad 0 \le u \le 1$$

son parametrizaciones equivalentes del cuarto de circunferencia.
(c) Encuentre dos parametrizaciones no equivalentes de alguna curva.

4. Demuestre que la curva

$$(x, y) = (\cos s, \mathrm{sen}\,s), \qquad 0 \le s \le 2\pi$$

está parametrizada por el largo de arco y dibuje los vectores velocidad y aceleración en $s = \pi/2$.

5. Sea γ la curva continuamente diferenciable de extremos $\mathbf{p}_1$ y $\mathbf{p}_2$. Sea λ el segmento de rectas $\mathbf{p}_1 + t(\mathbf{p}_2 - \mathbf{p}_1)$, $0 \le t \le 1$. Demuestre que $l(\lambda) \le l(\gamma)$.

6. Considere la curva espiral

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a\cos\omega t \\ a\,\mathrm{sen}\,\omega t \\ bt \end{pmatrix}, \qquad 0 \le t.$$

(a) Encuentre explícitamente la parametrización por largo de arco de esta curva.
(b) Encuentre el vector tangente unitario y el vector normal principal en un punto arbitrario.
(c) Encuentre la curvatura en un punto arbitrario.

7. (a) Demuestre que la recta definida por $g(t) = t\mathbf{x}_1 + \mathbf{x}_0$ tiene curvatura idénticamente cero.
(b) Demuestre que una curva γ parametrizada por el largo de arco y definida por la función $f(s)$ que tiene una tangente en cada punto y curvatura idénticamente igual a cero, es una recta.

8. Encuentre la masa total de la espiral definida por $g(t) = (a\cos t,\ a\,\mathrm{sen}\,t,\ bt)$, $0 \le t \le 2\pi$, si su densidad por unidad de largo en (x, y, z) es $x^2 + y^2 + z^2$.

9. Demuestre que si γ es el *gráfico* de la función $\mathcal{R} \xrightarrow{f} \mathcal{R}^2$ definida para $a \le x \le b$, entonces

$$l(\gamma) = \int_a^b \sqrt{1 + (f_1'(x))^2 + (f_2'(x))^2}\; dx,$$

donde f_1 y f_2 son las funciones coordenadas de γ que se supone son continuamente diferenciables.

10. Si una curva está parametrizada por el largo de arco, su curvatura es $\theta(s) = |(d/ds)\mathbf{t}(s)|$. Demuestre que si $\theta(s, h)$ es el ángulo entre $\mathbf{t}(s)$ y $\mathbf{t}(s + h)$, entonces

$$\kappa(s) = \lim_{h\to 0}\left|\frac{\theta(s, h)}{h}\right|.$$

[*Indicación*: Demuestre que $|\mathbf{t}(s + h) - \mathbf{t}(s)| = \sqrt{2 - 2\cos\theta(s, h)}$.]

11. Demuestre que si una curva está dada paramétricamente por la función $g(t)$, entonces la curvatura en $g(t)$ se expresa en términos de sus derivadas con respecto a t, por

$$\frac{\sqrt{|g'(t)|^2|g''(t)|^2 - (g'(t)\cdot g''(t))^2}}{|g'(t)|^3}.$$

[*Indicación*: Exprese $|\mathbf{t}'|$ en términos de g' y g''.]

12. (a) Encuentre la función curvatura $\kappa(t)$ de la parábola $\mathbf{x}(t) = (t, t^2)$ para $-\infty < t < \infty$.

(b) Demuestre que la curvatura de un círculo de radio r está dada por $\kappa(t) = 1/r$.

13. Sea γ una curva continuamente diferenciable que tiene una densidad de distribución de masa $p(\mathbf{x})$ en cada punto $\mathbf{x}$ de γ. Sea m la masa total de γ. Si γ está definida por $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, $a \le t \le b$, el **centro de masa** de la distribución es el vector

$$\mathbf{z} = \frac{1}{m}\int_a^b g(t)p(g(t))\,|g'(t)|\,dt.$$

(Véase el Ejercicio 14 de la sección anterior para la definición de la integral vectorial.)

(a) Encuentre el centro de masa de la espiral $g(t) = (a\cos t, a\,\text{sen}\,t, bt)$, $0 \le t \le 2\pi$, con densidad en (x, y, z) igual a $x^2 + y^2 + z^2$.

(b) Demuestre que si γ tiene densidad uniforme 1 y está parametrizada por el largo de arco, entonces

$$\mathbf{z} = \frac{1}{m}\int_0^{l(\gamma)} g(s)\,ds.$$

(c) Use el resultado del Problema 17(b) de la sección anterior para demostrar que el centro de masa satisface

$$|\mathbf{z}| \le \frac{1}{m}\int_0^{l(\gamma)} |g(s)|\,ds.$$

4. Integrales de línea

La integral $\int_a^b f(x)\,dx$ de una función real f de una variable real puede ser generalizada de varias maneras. Una generalización que tiene aplicación en la física es la integral de línea que describiremos aquí. Sea $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}^3$ un campo vectorial continuo definido en la región D de $\mathcal{R}^3$. Sea γ una curva en D parametrizada por una función $\mathcal{R} \xrightarrow{g} \mathcal{R}^3$ con $g(t)$ definida y continuamente diferenciable para $a \le t \le b$. Decir que γ está en D, significa simplemente que la imagen de g está contenida en D. Una situación típica es la de la Fig. 19. En cada punto $g(t)$ de γ nos imaginamos al vector $g'(t)$ como una flecha cuyo punto inicial es $g(t)$. Además, ubicamos en $g(t)$ el vector que describe al campo vectorial F en $g(t)$. El producto punto $F(g(t)) \cdot g'(t)$ es entonces una función real continua de t para $a \le t \le b$ y la integral

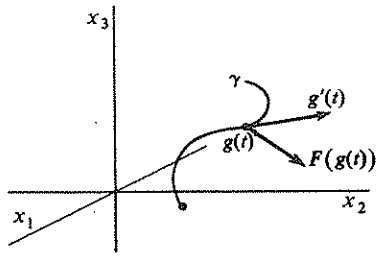


Figura 19

$$\int_a^b F(g(t)) \cdot g'(t)\,dt$$

se llama la **integral de línea** de F a lo largo de γ.

Ejemplo 1. Si un campo vectorial está dado en $\mathcal{R}^3$ por $F(x, y, z) = (x^2, y^2, z^2)$ y γ está definida por $g(t) = (t, t^2, t^3)$ para $0 \le t \le 1$, entonces la integral de F a lo largo de γ es

$$\int_0^1 (t^2, t^4, t^6) \cdot (1, 2t, 3t^2)\,dt = \int_0^1 (t^2 + 2t^5 + 3t^8)\,dt = 1.$$

La integral de línea puede ser interpretada en forma cualitativa como sigue. El producto punto

$$F(g(t)) \cdot \frac{g'(t)}{|g'(t)|}$$

es la coordenada de $F(g(t))$ en la dirección del vector unitario tangente a γ en $g(t)$. Entonces el integrando $F(g(t)) \cdot g'(t)$ de la Ec. (1), es la coordenada tangencial de $F(g(t))$ multiplicada por la rapidez $|g'(t)|$ con que se recorre a γ en $g(t)$. En particular, si $F(g(t))$ es siempre perpendicular a γ en $g(t)$, el integrando y, por tanto, la integral, es cero. En el otro extremo, dado un campo F, y si la rapidez $|g'(t)|$ está prescrita en cada punto de la curva, la integral será maximizada si se elige una curva γ cuya dirección en cada punto es la misma que la dirección de F en ese punto. Por tanto, el integrando de la integral de línea puede ser interpretada como una medida del efecto del campo vectorial a lo largo de γ.

La Ec. (1) puede ser generalizada a cualquier número de dimensiones. Así, si $\mathcal{R}^n \xrightarrow{F} \mathcal{R}^n$ y $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, $a \le t \le b$ es una curva lisa γ en D, la integral de línea de F a lo largo de γ está definida por la Ec. (1), en la cual el producto punto se toma en $\mathcal{R}^n$.

Ejemplo 2. Sea $F(x, y) = (x, y)$ un campo vectorial en $\mathcal{R}^2$. La curva

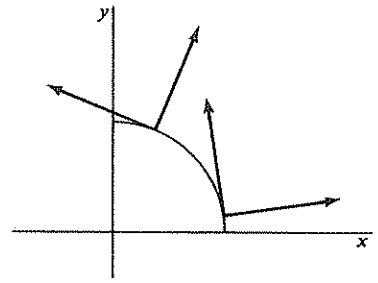


Figura 20

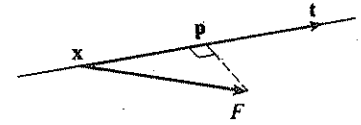


Figura 21

definida por $g(t) = (\cos t, \operatorname{sen} t)$ para $0 \le t \le \pi/2$ es el cuarto de circunferencia mostrado en la Fig. 20, donde también se muestran algunos vectores tangentes y algunos vectores del campo. Como el campo es perpendicular a la curva en cada punto, esperamos que la integral sea cero y, en efecto, tenemos

$$\int_0^{\pi/2} F(g(t)) \cdot g'(t)\, dt = \int_0^{\pi/2} (\cos t, \operatorname{sen} t) \cdot (-\operatorname{sen} t, \cos t)\, dt$$
$$= \int_0^{\pi/2} (-\cos t \operatorname{sen} t + \operatorname{sen} t \cos t)\, dt = 0.$$

Ejemplo 3. Una interpretación física importante de la integral de línea es la siguiente: Supongamos que la función $\mathbb{R}^3 \xrightarrow{F} \mathbb{R}^3$ determina un campo continuo de fuerzas en una región D de $\mathbb{R}^3$. Para definir el trabajo W efectuado al mover una partícula a lo largo de γ en D, usamos la definición que dice que, para un movimiento rectilíneo en un campo constante, el trabajo es fuerza multiplicada por distancia. Es decir,

$$W = (F_{\mathbf{t}})(s),$$

donde s es la distancia recorrida y $F_{\mathbf{t}}$ es la coordenada de la fuerza en la dirección del movimiento. En la Fig. 21, una partícula se mueve en una recta que tiene la dirección del vector unitario $\mathbf{t}$ y en cada punto $\mathbf{x}$ de ella actúa el mismo vector fuerza $F(\mathbf{x})$. La coordenada de F en la dirección del movimiento es $F_{\mathbf{t}} = F \cdot \mathbf{t}$. Por tanto, $W = (F \cdot \mathbf{t})s$.

En el caso de un movimiento a lo largo de una curva continuamente diferenciable, empezamos por aproximar la curva por vectores tangentes, como se hiciera al definir el largo de arco. Si la curva está definida paramétricamente por $\mathbb{R} \xrightarrow{g} \mathbb{R}^3$, donde $g(t)$ está definido para $a \le t \le b$, entonces las flechas que representan a los vectores tangentes $g'(t_{k-1})(t_k - t_{k-1})$, $t_0 < t_1 < \ldots < t_K$, forman una aproximación de γ como la mostrada en la Fig. 22. Fijemos un punto $\mathbf{x}_k = g(t_k)$ en γ y cerca de $\mathbf{x}_k$ aproximamos $F(\mathbf{x})$ por el campo vectorial que asocia a cada punto el vector constante $F(\mathbf{x}_k)$. La coordenada tangencial de $F(\mathbf{x}_k)$ es $F(\mathbf{x}_k) \cdot \mathbf{t}(t_k)$, donde $\mathbf{t}(t) = g'(t)/|g'(t)|$. Por tanto, el trabajo efectuado al mover la partícula desde $\mathbf{x}_k$ a $\mathbf{x}_{k+1}$ a lo largo de γ es, aproximadamente,

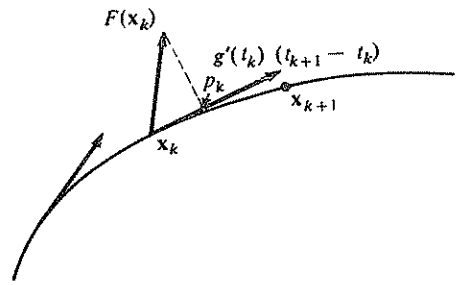


Figura 22

$$W_k = (F(\mathbf{x}_k) \cdot \mathbf{t}(t_k))\,|g'(t_k)|\,(t_{k+1} - t_k)$$
$$= F(g(t_k)) \cdot g'(t_k)(t_{k+1} - t_k).$$

Sea $m(P) = \max_{1 \le k \le K} (t_k - t_{k-1})$, entonces

$$\lim_{m(P) \to 0} \sum_{k=0}^{K-1} W_k = \int_a^b F(g(t)) \cdot g'(t)\, dt,$$

que definimos como el **trabajo** efectuado al mover la partícula en el dominio de F a lo largo de γ.

Las hipótesis hechas en el Ejemplo 3 de que F y g' son continuas, aseguran que el integrando $F(g(t)) \cdot g'(t)$ es continuo y que, por tanto, la integral existe. Sin embargo, estas hipótesis son más restrictivas que lo necesario. Es suficiente suponer que el camino de integración es liso por pedazos y que el campo vectorial F es lo suficientemente regular como para que la integral de la Ec. (1) exista. Es así cómo g' puede ser discontinua en un número finito de puntos, de modo que γ tenga esquinas como las mostradas en la Fig. 23.

Ejemplo 4. Consideremos el campo vectorial en $\mathbb{R}^3$, definido por $F(x, y, z) = (x, y, z)$. Sea $g(t) = (\cos t, \operatorname{sen} t, |t - \pi/2|)$ la descripción de una curva γ en $\mathbb{R}^3$ para $0 \le t \le \pi$. Entonces γ tiene una esquina en $(0, 1, 0)$, donde $t = \pi/2$. g no es diferenciable en este punto y, además, se tiene que $\lim_{t \to \pi/2-} g'(t) = (-1, 0, -1)$ y $\lim_{t \to \pi/2+} g'(t) = (-1, 0, 1)$, lo cual muestra que la dirección de las tangentes da un salto abrupto en $t = \pi/2$. Sin embargo, la integral de F a lo largo de γ existe. Para calcularla, se puede dividir el intervalo de integración en dos partes mediante el punto $t = \pi/2$. Sin embargo, en este caso especial, $F(g(t)) \cdot g'(t) = t - \pi/2$ para $0 \le t \le \pi$, de modo que, como se ve fácilmente, la integral de línea es cero.

Si escribimos la parametrización de γ en la forma $g(t) = (x(t), y(t), z(t))$ para $a \le t \le b$, podemos introducir una notación muy conveniente para las

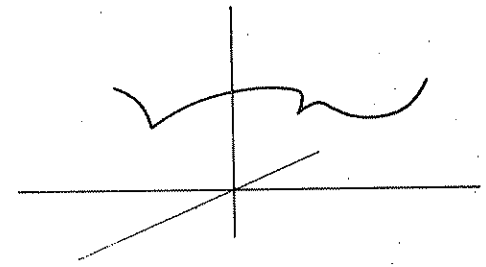


Figura 23

integrales de línea. Si denotamos por F_1, F_2 y F_3 las funciones coordenadas de F y suprimimos la variable t en el integrando, tenemos

$$\int_a^b F(g(t)) \cdot g'(t)\, dt = \int_a^b \left[F_1(x, y, z) \frac{dx}{dt} + F_2(x, y, z) \frac{dy}{dt} + F_3(x, y, z) \frac{dz}{dt} \right] dt.$$

Esta última integral puede ser abreviada en la forma

$$\int_\gamma F_1\, dx + F_2\, dy + F_3\, dz$$

la cual puede ser abreviada aún más si escribimos $d\mathbf{x} = (dx, dy, dz)$. Entonces la fórmula se transforma en

$$\int_\gamma F \cdot d\mathbf{x}.$$

En el Cap. 7, Sec. 7, veremos en mayor detalle el significado de esta notación. Por ahora la consideraremos sencillamente como una abreviación.

Ejemplo 5. Sea $\mathcal{R}^4 \xrightarrow{F} \mathcal{R}^4$ definida por

$$F(x, y, z, w) = (x - y, y - z, z - w, w - x).$$

La curva γ definida por $g(t) = t, -t, t^2, -t^2)$ para $0 \leq t \leq 1$ pasa por la región de definición del campo. La integral de F a lo largo de γ es

$$\int_\gamma F \cdot d\mathbf{x} = \int_0^1 [(2t)(1) + (-t - t^2)(-1) + (2t^2)(2t) + (-t^2 - t)(-2t)]\, dt = 4.$$

Sea γ una curva diferenciable en $\mathcal{R}^n$ definida por $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ para $a \leq t \leq b$. La integral a lo largo de γ de una función real p cuyo dominio contiene a γ se define como

$$\int_a^b p(g(t))\, |g'(t)|\, dt. \tag{2}$$

La función p puede ser interpretada como la densidad por unidad de longitud de una distribución de masa en γ en cuyo caso la integral representa la masa total distribuida a lo largo de γ. Si escribimos la integral de línea de un campo vectorial a lo largo de γ en la forma

$$\int_a^b F(g(t)) \cdot \frac{g'(t)}{|g'(t)|}\, |g'(t)|\, dt, \tag{3}$$

entonces queda clara la relación entre las Ecs. (2) y (3). La integral de un campo vectorial a lo largo de γ depende de la dirección de la tangente de γ en cada punto y no solo de su largo como es el caso en la integral de una función

real a lo largo de γ. Si la curva γ está parametrizada por el largo de arco, entonces las dos integrales toman las formas

$$\int_0^{u(\gamma)} p(g(s))\, ds \quad \text{y} \quad \int_0^{u(\gamma)} F(g(s)) \cdot \mathbf{t}(s)\, ds,$$

donde $\mathbf{t}(s) = (dg/ds)(s)$ es un vector tangente a γ de largo 1.

Es claro que, por definición, el valor de la integral de línea depende de la parametrización de la curva γ. En los ejercicios se verá hasta qué punto este valor es independiente de la parametrización.

EJERCICIOS

1. Calcule las siguientes integrales de línea:
   (a) $\int_L x\, dx + x^2\, dy + y\, dz$, donde L está dada por $g(t) = (t, t, t)$, para $0 \leq t \leq 1$.
   (b) $\int_P (x + y)\, dx + dy$, donde P está dada por $g(t) = (t, t^2)$, $0 \leq t \leq 1$.
   (c) $\int_{\gamma_1} x\, dy$ y $\int_{\gamma_2} x\, dy$, donde γ_1 está definida por $g(t) = (\cos t, \operatorname{sen} t)$ para $0 \leq t \leq 2\pi$ y γ_2 está definida por $h(t) = (\cos t, \operatorname{sen} t)$ para $0 \leq t \leq 4\pi$.
   (d) $\int_{\gamma_1} (dx + dy)$, donde γ_1 está definida por $(x, y) = (\cos t, \operatorname{sen} t)$, $0 \leq t \leq 2$.
   (e) $\int_{\gamma_1} \frac{dx + dy}{x^2 + y^2}$, donde γ_1 es la curva de la parte (d).
   (f) $\int_\gamma (e^x\, dx + z\, dy + \operatorname{sen} z\, dz)$, donde γ está definida por $(x, y, z) = (t, t^2, t^3)$, $0 \leq t \leq 1$.
   (g) $\int_\gamma F \cdot d\mathbf{x}$, donde $F(x, y, z, w) = (x, x, y, xw)$ y γ está definida por $(x, y, z, w) = (t, 1, t, t)$, $0 \leq t \leq 2$.

2. Supongamos que γ_1 está definida por $(x, y) = (\cos t, \operatorname{sen} t)$, $0 \leq t \leq 2\pi$ y γ_2 por $(x, y) = (1 - u, u)$, $0 \leq u \leq 1$. Calcule $\int_{\gamma_1} (f\, dx + g\, dy)$ y $\int_{\gamma_2} (f\, dx + g\, dy)$ para los siguientes f, g:
   (a) $f(x, y) = x, g(x, y) = x + 1$.
   (b) $f(x, y) = x + y, g(x, y) = 1$.
   (c) $f(x, y) = \frac{1}{x^2 + y^2}, g(x, y) = \frac{1}{x^2 + y^2}$.

3. Calcule el trabajo efectuado al mover una partícula a lo largo de la curva $(x, y, z) = (t, t, t^2)$, $0 \leq t \leq 2$, bajo la influencia del campo $F(x, y, z) = (x + y, y, y)$.

4. Demuestre que si la curva γ está definida paramétricamente por una función $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$ donde
$$\mathbf{x} = f(t), \qquad 0 \leq t \leq 1,$$
y si $-\gamma$ es la curva definida por
$$\mathbf{x} = f(1 - t), \qquad 0 \leq t \leq 1,$$
entonces
$$\int_{-\gamma} F \cdot d\mathbf{x} = -\int_\gamma F \cdot d\mathbf{x}.$$

5. Sea $F(x, y) = (y, x)$ un campo vectorial en $\mathcal{R}^2$. Encuentre una curva γ_1 que pase

por el campo, partiendo del punto (2, 1) tal que γ_1 tenga largo 1 y que la integral de F a lo largo de γ_1 sea cero.

6. Demuestre que si γ está definida por $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$, $a \le t \le b$, entonces si se reparametriza γ por el largo de arco, de modo que $t = t(s)$, se tiene que

$$\int_a^b F(g(t)) \cdot g'(t)\, dt = \int_0^{t(\gamma)} F(h(s)) \cdot \mathbf{t}(s)\, ds,$$

donde $h(s) = g(t(s))$ y $\mathbf{t}(s) = (dh/ds)(s)$. [*Indicación*: Use el teorema sobre el cambio de variables en una integral.]

7. Sea $\mathcal{R}^n \xrightarrow{F} \mathcal{R}^n$ un campo vectorial y γ una curva tal que la integral de línea $\int_\gamma F \cdot d\mathbf{x}$ existe. Demuestre que si se tiene que, en γ $|F(\mathbf{x})| \le M$, donde M es una constante, entonces $\left|\int_\gamma F \cdot d\mathbf{x}\right| \le M l(\gamma)$.

8. (a) Si $g(t)$ y $h(u)$ son parametrizaciones equivalentes de γ (véase definición de equivalencia en el Ejercicio 3 de la Sec. 4), demuestre que si $\int_\gamma F \cdot d\mathbf{x}$ se calcula usando la parametrización $g \circ h$, se obtiene el mismo resultado.
   (b) Encuentre parametrizaciones no equivalentes de la circunferencia $x^2 + y^2 = 1$ en $\mathcal{R}^2$ y un campo vectorial F tal que las integrales de F con respecto a las dos parametrizaciones son diferentes.

9. Encuentre la masa total del alambre $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6t^2 \\ 4\sqrt{2}\, t^3 \\ 3t^4 \end{pmatrix}$, $0 \le t \le 1$:
   (a) Si la densidad en el punto que corresponde a t es t^2.
   (b) Si la densidad en un punto a una distancia s del origen a lo largo de la curva es $s + 1$.
   (c) Si la densidad en un punto es igual a su distancia al origen, medida en $\mathcal{R}^3$.

10. Demuestre que si $\int_\gamma F \cdot d\mathbf{x}$ y $\int_\gamma G \cdot d\mathbf{x}$ existen, entonces $\int_\gamma (aF + bG) \cdot d\mathbf{x} = a \int_\gamma F \cdot d\mathbf{x} + b \int_\gamma G \cdot d\mathbf{x}$, donde a y b son dos constantes cualesquiera.

11. Demuestre que si γ y η son dos curvas lisas descritas por las funciones $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ definidas en $[a, b]$ y $\mathcal{R} \xrightarrow{h} \mathcal{R}^n$ definida en $[b, c]$ y tales que $g(b) = h(b)$, entonces

$$\int_\delta F \cdot d\mathbf{x} = \int_\gamma F \cdot d\mathbf{x} + \int_\eta F \cdot d\mathbf{x},$$

donde δ es la curva definida por

$$f(t) = \begin{cases} g(t), & a \le t \le b \\ h(t), & b \le t \le c, \end{cases}$$

y F es un campo vectorial continuo en $\gamma \cup \eta$.

12. Supongamos que la función $g(t)$ representa la posición en $\mathcal{R}^3$, de una partícula de masa variable $m(t)$, en el tiempo t. Entonces el vector velocidad de la partícula es $\mathbf{v}(t) = g'(t)$ y la fuerza que actúa sobre la partícula en $g(t)$ es $F(g(t)) = [m(t)\mathbf{v}(t)]'$.
   (a) Demuestre que $F(g(t)) \cdot g'(t) = m(t)v^2(t) + m(t)v(t)v'(t)$, donde v es la rapidez de la partícula.

   (b) Demuestre que si $m(t)$ es constante, entonces el trabajo efectuado al mover la partícula sobre su camino, entre los tiempos $t = a$ y $t = b$, es $w = (m/2)(v^2(b) - v^2(a))$. (La función $(\frac{1}{2})$ en $v^2(t)$ es la energía cinética de la partícula.)

5. Derivadas parciales

Para extender las técnicas del cálculo a funciones definidas en $\mathcal{R}^n$, necesitamos las derivadas parciales. Sea f una función real con espacio dominio $\mathcal{R}^n$. Para cada $i = 1, \ldots, n$, definimos una nueva función real llamada la **derivada parcial** de f **con respecto de la i-ésima variable** y que denotamos con $\partial f/\partial x_i$. Para cada $\mathbf{x} = (x_1, \ldots, x_n)$ en el dominio de f, el número $(\partial f/\partial x_i)(\mathbf{x})$ es, por definición,

$$\mathbf{5.1} \qquad \frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{t \to 0} \frac{f(x_1, \ldots, x_i + t, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{t}.$$

El espacio dominio de $\partial f/dx_i$ es $\mathcal{R}^n$ y su dominio es un subconjunto del dominio de f y está formado por todos los $\mathbf{x}$ para los cuales el límite indicado existe. Por tanto, este dominio podría ser vacío. El número $(\partial f/\partial x_i)(\mathbf{x})$ es simplemente la derivada de la función de una variable obtenida manteniendo fijos $x_1, \ldots, x_{i-1}$, $x_{i+1}, \ldots, x_n$ y considerando f como función solo de la i-ésima variable.

Ejemplo 1. Sea $f(x, y, z) = x^2y + y^2z + z^2x$. Entonces

$$\frac{\partial f}{\partial x}(x, y, z) = 2xy + z^2,$$

$$\frac{\partial f}{\partial y}(x, y, z) = x^2 + 2yz,$$

$$\frac{\partial f}{\partial z}(x, y, z) = y^2 + 2zx.$$

Las derivadas parciales en $\mathbf{x} = (1, 2, 3)$ son

$$\frac{\partial f}{\partial x}(1, 2, 3) = 4 + 9 = 13,$$

$$\frac{\partial f}{\partial y}(1, 2, 3) = 1 + 12 = 13,$$

$$\frac{\partial f}{\partial z}(1, 2, 3) = 4 + 6 = 10.$$

Ejemplo 2. Sea $f(u, v) = \operatorname{sen} u \cos v$. Entonces

$$\frac{\partial f}{\partial u} = \frac{\partial \operatorname{sen} u \cos v}{\partial u} = \cos u \cos v,$$

$$\frac{\partial f}{\partial v} = \frac{\partial \operatorname{sen} u \cos v}{\partial v} = -\operatorname{sen} u \operatorname{sen} v,$$

$$\frac{\partial f}{\partial v}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = \frac{\partial \operatorname{sen} u \cos v}{\partial v}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = -\operatorname{sen}\frac{\pi}{2} \operatorname{sen}\frac{\pi}{2} = -1.$$

La operación de tomar derivadas parciales puede ser repetida. La derivada parcial de $\partial f/\partial x_i$ con respecto a la j-ésima variable es $\partial/\partial x_j\,(\partial f/\partial x_i)$ y se denota por $\partial^2 f/\partial x_j\,\partial x_i$. Esto puede ser repetido indefinidamente, siempre que las derivadas existan. Una notación distinta para las derivadas parciales de orden superior se ilustra más adelante, en la cual la variable de derivación se denota por un subíndice.

$$\frac{\partial f}{\partial x_i} = f_{x_i}$$

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j\,\partial x_i} = f_{x_i x_j}$$

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_i^2} = f_{x_i x_i}$$

$$\frac{\partial}{\partial x_i}\left(\frac{\partial^2 f}{\partial x_j\,\partial x_i}\right) = \frac{\partial^3 f}{\partial x_i\,\partial x_j\,\partial x_i} = f_{x_i x_j x_i}.$$

Ejemplo 3. Consideremos $f(x, y) = xy - x^2$.

$$f_x = \frac{\partial f}{\partial x} = y - 2x$$

$$f_{xy} = \frac{\partial^2 f}{\partial y\,\partial x} = 1$$

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = -2$$

$$f_{yxx} = \frac{\partial^3 f}{\partial x^2\,\partial y} = 0.$$

Para interpretar geométricamente las derivadas parciales podemos usar el hecho de que para una función real de una variable, el valor de la derivada en un punto es la pendiente de la recta tangente al gráfico de la función en el punto considerado. Como ilustración, consideraremos el gráfico de una función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$, es decir, el conjunto de punto $(x, y, f(x, y))$ en $\mathcal{R}^3$, donde (x, y) está en el dominio de f. Un ejemplo se muestra en la Fig. 24. La intersección de la superficie con el plano $y = b$ es una curva, definida implícitamente por las ecuaciones

$$z = f(x, y), \qquad y = b.$$

Consideremos la curva definida por la función $g(x) = f(x, b)$ como una curva en el espacio de dos dimensiones. Su pendiente en $x = a$ es

$$g'(a) = \frac{\partial f}{\partial x}(a, b).$$

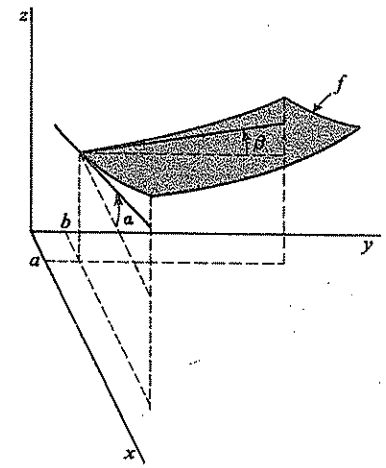


Figura 24

En forma análoga, la curva definida por $h(y) = f(a, y)$ tiene pendiente en $y = b$ igual a

$$h'(b) = \frac{\partial f}{\partial y}(a, b).$$

Entonces, los ángulos α y β mostrados en la Fig. 24 satisfacen

$$\tan\alpha = \frac{\partial f}{\partial x}(a, b), \qquad \tan\beta = \frac{\partial f}{\partial y}(a, b).$$

Los números $\tan\alpha$ y $\tan\beta$ son las pendientes de las rectas tangentes a dos curvas contenidas en el gráfico de f. Por esta razón es natural definir el plano tangente al gráfico de f como el plano que contiene a estas dos rectas tangentes. (Si f satisface la condición de diferenciabilidad que se dará en la Sec. 8, intro-

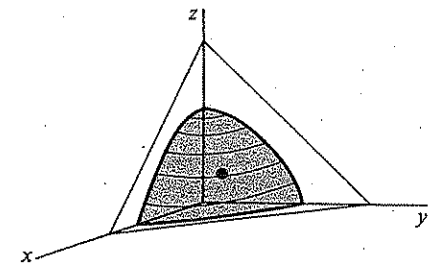


Figura 25

duciremos en efecto esta definición.) Se ve fácilmente que el conjunto de los puntos (x, y, z) que satisfacen la ecuación

$$z = f(a, b) + (x - a)\frac{\partial f}{\partial x}(a, b) + (y - b)\frac{\partial f}{\partial y}(a, b) \qquad (1)$$

es un plano que contiene las rectas tangentes encontradas anteriormente. En efecto, si $y = b$ o $x = z$, obtenemos esas rectas.

Ejemplo 4. La parte del gráfico de la función

$$f(x, y) = 1 - 2x^2 - y^2$$

que corresponde a $x \geq 0$, $y \geq 0$ es mostrada en la Fig. 25. Las derivadas parciales de f en $(\frac{1}{2}, \frac{1}{2})$ son

$$\frac{\partial f}{\partial x}\left(\frac{1}{2}, \frac{1}{2}\right) = -2, \qquad \frac{\partial f}{\partial y}\left(\frac{1}{2}, \frac{1}{2}\right) = -1.$$

Usando la Ec. (1) para definir el plano tangente al gráfico de f en $(\frac{1}{2}, \frac{1}{2})$, y puesto que $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$, tenemos

$$\begin{aligned} z &= \tfrac{1}{4} - 2(x - \tfrac{1}{2}) - (y - \tfrac{1}{2}) \\ &= -2x - y + \tfrac{7}{4}. \end{aligned}$$

Podemos dibujar el plano tangente trazando las dos rectas tangentes en él determinadas por $x = \frac{1}{2}$ e $y = \frac{1}{2}$. Es más fácil ubicar tres puntos en el plano, por ejemplo,

$$(\tfrac{7}{8}, 0, 0), \qquad (0, \tfrac{7}{4}, 0), \qquad (0, 0, \tfrac{7}{4}).$$

El punto de tangencia es $(\frac{1}{2}, \frac{1}{2}, \frac{1}{4})$.

En la Sec. 7 se estudiará la continuidad de funciones de más de una variable. En este lugar consideraremos brevemente el caso de una función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$. Por conveniencia, suponemos que, para todo $\mathbf{x} = (x, y)$ en el dominio de f, $f(\mathbf{z})$ está definido para todos los vectores $\mathbf{z} = (z, w)$, tales que $|\mathbf{x} - \mathbf{z}| < \delta$, donde δ es algún número positivo. Decimos entonces que f es continua si, para todo punto $\mathbf{x}$ en el dominio de f,

$$\lim_{|\mathbf{x}-\mathbf{z}|\to 0} f(\mathbf{z}) = f(\mathbf{x}).$$

Esta relación significa que si la distancia $|\mathbf{x} - \mathbf{z}|$ entre $\mathbf{x}$ y $\mathbf{z}$ se hace suficientemente pequeña, $f(\mathbf{z})$ se acerca tanto como se desea a $f(\mathbf{x})$. Como de costumbre, la idea intuitiva de continuidad es que los valores de la función f no cambian abruptamente, es decir, que no hay rupturas en el gráfico de f. Los gráficos de las Figs. 24 y 25 representan funciones continuas, mientras que la Fig. 26 muestra un ejemplo de una función discontinua. La condición de que el dominio D de f contenga todos los puntos $\mathbf{z}$ suficientemente cercanos a cada $\mathbf{x}$ en D, se expresa diciendo que D es un **conjunto abierto**.

Una consecuencia de ciertas condiciones de continuidad de f es que las derivadas parciales de orden superior de $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$ son independientes del

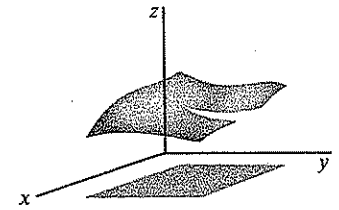


Figura 26

orden de derivación. A continuación damos un enunciado preciso, haciendo notar que también se puede demostrar un resultado algo más fuerte. (Véase el Ejercicio 8 de la Sec. 2 del Cap. 6.)

5.2 Teorema

Sea $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$ una función continua tal que f_x, f_y, f_{xy} y f_{yx} son continuas en el de f. Entonces $f_{xy} = f_{yx}$.

Demostración. Elijamos x, y y $\delta > 0$, de modo que la función

$$\begin{aligned} F(h, k) = &[f(x + h, y + k) - f(x + h, y)] \\ &- [f(x, y + k) - f(x, y)] \end{aligned}$$

esté definida siempre que $\sqrt{h^2 + k^2} < \delta$. Supongamos que h y k no son cero y apliquemos el teorema del valor medio a la función

$$G(u) = f(u, y + k) - f(u, y)$$

en el intervalo con extremos x y $x + h$. Obtenemos

$$G(x + h) - G(x) = hG'(x_1),$$

donde x_1 está entre x y $x + h$. En términos de F y f, esta última ecuación toma la forma

$$F(h, k) = h[f_x(x_1, y + k) - f_x(x_1, y)].$$

Aplicamos el teorema del valor medio a la función $H(v) = f_x(x_1, v)$ en el intervalo con extremos y e $y + k$, obteniendo

$$F(h, k) = hkf_{xy}(x_1, y_1),$$

donde y_1 está entre y e $y + k$.

Reescribiendo F en la forma

$$\begin{aligned} F(h, k) = &[f(x + h, y + k) - f(x, y + k)] \\ &- [f(x + h, y) - f(x, y)] \end{aligned}$$

podemos aplicar el mismo procedimiento general, diferenciando primero con respecto de y y luego respecto de x. Obtenemos

$$F(h, k) = hkf_{yx}(x_2, y_2),$$

donde x_2 está en $(x, x + h)$ e y_2 en $(y, y + k)$. Igualmente, las dos expresiones para $F(h, k)$ y cancelando el factor hk, tenemos

$$f_{xy}(x_1, y_1) = f_{yx}(x_2, y_2).$$

Si hacemos tender h y k a cero, las distancias

$$\sqrt{(x_1 - x)^2 + (y_1 - y)^2} \quad \text{y} \quad \sqrt{(x_2 - x)^2 + (y_2 - y)^2}$$

tienden a cero y, por continuidad de f_{xy} y f_{yx}, obtenemos que $f_{xy}(x, y) = f_{yx}(x, y)$. Como (x, y) es un arbitrario, $f_{xy} = f_{yx}$ en el dominio de f.

El Teorema 5.2 puede ser aplicado sucesivamente a derivadas de orden superior, siempre que se satisfagan condiciones análogas de diferenciabilidad y continuidad. Además, considerando las variables de dos en dos, lo podemos aplicar a funciones $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ con $n > 2$. Es así cómo para las funciones que se encuentran generalmente y que tienen derivadas parciales de todo orden, se tienen relaciones tales como

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 f}{\partial y\, \partial x}$$

$$\frac{\partial^3 g}{\partial x\, \partial y\, \partial x} = \frac{\partial^3 g}{\partial x^2\, \partial y}$$

$$\frac{\partial^4 h}{\partial z\, \partial x\, \partial y\, \partial z} = \frac{\partial^4 h}{\partial x\, \partial y\, \partial z^2}, \quad \text{etc.}$$

Las dos últimas fórmulas se obtienen aplicando repetidamente la fórmula para dos variables, intercambiando las diferenciaciones de dos en dos.

1. Encuentre $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ si $f(x, y)$ es

   (a) $x^2 + x \operatorname{sen}(x + y)$.
   (b) $\operatorname{sen} x \cos(x + y)$.
   (c) e^{x+y+1}.
   (d) $\arctan(y/x)$.
   (e) x^y.
   (f) $\log_x y$.

   [*Resp*: $\frac{\partial f}{\partial x} = -\frac{\ln y}{x(\ln x)^2}$.]

2. Encuentre $\frac{\partial^2 f}{\partial y\, \partial x}$ y $\frac{\partial^2 f}{\partial x\, \partial y}$ si f es

   (a) $xy + x^2y^3$. (b) $\operatorname{sen}(x^2 + y^2)$. (c) $\frac{1}{x^2 + y^2}$.

3. Encuentre las derivadas parciales de primer orden de las siguientes funciones:

   (a) $f(x, y, z) = x^2 e^{x+y+z} \cos y$.

   (b) $f(x, y, z, w) = \frac{x^2 - y^2}{z^2 + w^2}$.

   (c) $f(x, y, z) = x^{(y^z)}$.

4. Encuentre $\frac{\partial^3 f(x, y)}{\partial x^2\, \partial y}$ si $f(x, y) = \log(x + y)$.

5. Demuestre que la ecuación $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$ se satisface para

   (a) $\log(x^2 + y^2)$. (b) $x^3 - 3xy^2$.

6. Si $f(x, y, z) = 1/(x^2 + y^2 + z^2)^{1/2}$, demuestre que

$$f_{xx} + f_{yy} + f_{zz} = 0.$$

7. Si $f(x_1, x_2, \ldots, x_n) = 1/(x_1^2 + x_2^2 + \ldots x_n^2)^{(n-2)/2}$, demuestre que

$$f_{x_1x_1} + f_{x_2x_2} + \ldots + f_{x_nx_n} = 0.$$

8. Demuestre directamente que si f es un polinomio, entonces

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 f}{\partial y\, \partial x}.$$

9. Si

$$f(x, y) = \begin{cases} 2xy\dfrac{x^2 - y^2}{x^2 + y^2}, & \text{para } x^2 + y^2 \neq 0 \\ 0, & \text{para } x = y = 0, \end{cases}$$

   demuestre que $f_{xy}(0, 0) = -2$ y $f_{yx}(0, 0) = 2$.

10. Para cada una de las funciones que siguen, encuentre una representación explícita para el plano tangente al gráfico en el punto indicado. Dibuje además el gráfico de la función con su plano tangente.

    (a) $\sqrt{1 - x^2 - y^2}$ a $(\frac{1}{2}, \frac{1}{2}, 1/\sqrt{2})$.
    (b) e^{x+y} a $(1, 2, e^3)$.
    (c) $\operatorname{sen}(x^2 + y^2)$ a $(0, 0, 0)$.

11. Encuentre una representación paramétrica para la recta perpendicular al plano tangente encontrado en el Ejercicio 10(a) y que pasa por el punto de tangencia.

6. Derivadas parciales vectoriales

En la sección anterior solo consideramos las derivadas parciales de funciones reales. Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es una función vectorial, entonces la Ec. 5.1 define $\partial f/\partial x_i$ como el límite en una variable real. La diferencia radica en que ahora el cociente y, por tanto, el límite, es un vector. Puesto que, como lo describimos en la Secc. 2, los límites vectoriales se calculan tomando el límite de cada función coordenada, se tiene de inmediato que

6.1 Si

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \cdot \\ \cdot \\ \cdot \\ f_m(\mathbf{x}) \end{pmatrix},$$

entonces

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_i}(\mathbf{x}) \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial f_m}{\partial x_i}(\mathbf{x}) \end{pmatrix}.$$

El significado geométrico de la derivada parcial vectorial es el siguiente. Si se dejan fijas todas las coordenadas excepto una, digamos x_i, y solo se permite que x_i varíe, entonces $f(\mathbf{x})$ describe una curva en $\mathfrak{R}^n$. De acuerdo con la definición dada en la Sec. 2, el vector $\partial f/\partial x_i(\mathbf{x})$ es un vector tangente a esa curva.

Ejemplo 1. Sea

$$f(u, v) = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \\ v \end{pmatrix}.$$

Entonces

$$\frac{\partial f}{\partial u}(u_0, v_0) = \begin{pmatrix} \cos v_0 \\ \operatorname{sen} v_0 \\ 0 \end{pmatrix}, \qquad \frac{\partial f}{\partial v}(u_0, v_0) = \begin{pmatrix} -u_0 \operatorname{sen} v_0 \\ u_0 \cos v_0 \\ 1 \end{pmatrix}.$$

En el Ejemplo 1, podemos restringir el vector variable (u, v), de modo que $v = v_0$ se deja fijo. Entonces $f(u, v_0)$ describe una curva a medida que u varía. Análogamente, $f(u_0, v)$ describe otra curva al variar v. Estas curvas se llaman las curvas parámetros en la imagen de f. Supondremos que las funciones coordenadas de f son continuas. De acuerdo con la definición de la Sec. 2, los vectores

$$\frac{\partial f}{\partial u}(u_0, v_0) \quad \text{y} \quad \frac{\partial f}{\partial v}(u_0, v_0)$$

son vectores tangentes a las curvas parámetros obtenidas al hacer variar u y v, respectivamente, en el punto en que ellas se cortan; es decir, $f(u_0, v_0)$. La

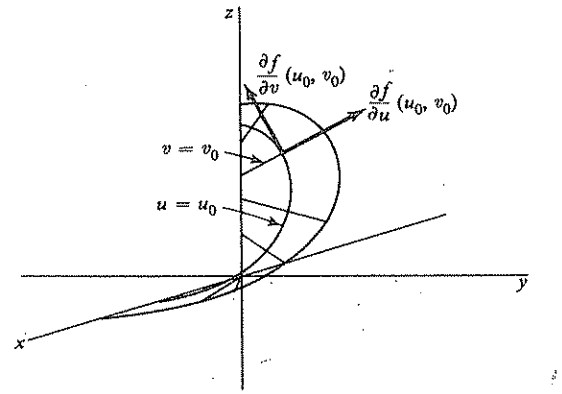


Figura 27

Fig. 27 muestra un ejemplo donde las dos curvas están dadas por las relaciones $v = v_0$ y $u = u_0$; el parámetro que varía en el caso $v = v_0$ es u y en el caso $u = u_0$ es v.

Bajo ciertas condiciones y, en particular en los ejemplos comunes, las curvas parámetros de una función $\mathfrak{R}^2 \xrightarrow{f} \mathfrak{R}^3$ se unen para formar una superficie en el espacio imagen $\mathfrak{R}^3$. En la última sección del Cap. 4 se discuten las condiciones técnicas para que ello suceda. En este lugar nos limitamos a considerar algunos ejemplos.

Ejemplo 2. Volviendo a la función f del Ejemplo 1, las curvas determinadas por una condición $u = u_0$ son espirales alrededor del eje vertical en $\mathfrak{R}^3$ a una distancia u_0 de él. En su conjunto, estas curvas forman una superficie con la forma de una rampa espiral tal como se viera en el Ejemplo 7 de la Sec. 1. También se puede pensar que la rampla está formada por las rectas determinadas por la condición $v = v_0$ que salen perpendicularmente del eje vertical. El punto $(u, v) = (1, \pi/2)$ en el dominio de f corresponde a

$$f(1, \pi/2) = (0, 1, \pi/2)$$

en la superficie. En este punto, los vectores tangentes

$$\frac{\partial f}{\partial u}(1, \pi/2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \frac{\partial f}{\partial v}(1, \pi/2) = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

determinan un plano que, al ser trasladado paralelamente de modo que pase por $f(1, \pi/2)$, forma un plano que es natural llamar un plano tangente. Una representación paramétrica de este plano tangente es

$$u\begin{pmatrix}0\\1\\0\end{pmatrix}+v\begin{pmatrix}-1\\0\\1\end{pmatrix}+\begin{pmatrix}0\\1\\\frac{\pi}{2}\end{pmatrix}.$$

En general, diremos que si la imagen de una función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^3$ es una superficie con un plano tangente en $f(u_0, v_0)$, entonces el plano tangente queda representado paramétricamente por

$$u\frac{\partial f}{\partial u}(u_0, v_0) + v\frac{\partial f}{\partial v}(u_0, v_0) + f(u_0, v_0),$$

con u y v como parámetros.

A veces una función $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}^2$ puede ser interpretada como un **flujo** en dos dimensiones de la manera siguiente: Un punto (x, y, t) en $\mathcal{R}^3$ es transformado por f en el punto $(x', y') = f(x, y, t)$ en $\mathcal{R}^2$, donde (x, y) se interpreta como la posición de una partícula en $\mathcal{R}^2$ en el instante $t = 0$ y $f(x, y, t)$ es la posición de la misma partícula después de un tiempo t. La Fig. 28 muestra un

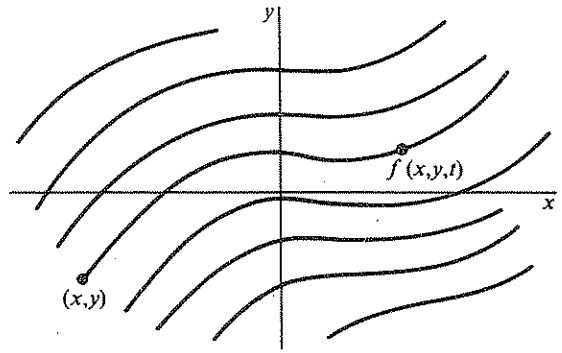


Figura 28

flujo en dos dimensiones. Allí, el camino de una partícula que está en la posición (x, y) en el instante $t = 0$ se muestra como una curva que contiene todos los puntos $f(x, y, t)$. Así, si mantenemos fijos (x, y) y hacemos variar t, obtenemos una curva llamada una **trayectoria** del flujo. La familia de todas las trayectorias determinadas por f da una representación geométrica del flujo. Los vectores tangentes

$$\frac{\partial f}{\partial t}(x, y, t),$$

para cada t fijo forman un campo vectorial en el cual la cola de la flecha $f_t(x, y, t)$ está en el punto $(x', y') = f(x, y, t)$ que inicialmente estaba en (x, y). Si este campo vectorial asociado a un flujo es independiente de t, el flujo se llama

estable. Como lo muestra el ejemplo que sigue, un flujo estable no es necesariamente uno en que $f_t(x, y, t)$ es independiente de t. Por otra parte, no se debe suponer que en un flujo no estable cada partícula en una trayectoria determinada sigue el mismo camino que inicia esa trayectoria.

Ejemplo 3. La función

$$f(x, y, t) = \begin{pmatrix}x\cos t - y\operatorname{sen} t\\ x\operatorname{sen} t + y\cos t\end{pmatrix} = \begin{pmatrix}\cos t & -\operatorname{sen} t\\ \operatorname{sen} t & \cos t\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

es una función lineal en (x, y) y es, para cada t, una rotación alrededor del origen de $\mathcal{R}^2$, en sentido contrario a el de los punteros del reloj en un ángulo t. Por tanto, la trayectoria de un punto (x, y) es un círculo de radio $\sqrt{x^2 + y^2}$ con centro en el origen. La Fig. 29 muestra alguna de las trayectoria. El vector velocidad $\partial f/\partial t(x, y, t)$ ha sido denotado por $f_t(x, y, t)$ en la Fig. 29. En particular,

$$f_t(x, y, t) = \begin{pmatrix}-x\operatorname{sen} t - y\cos t\\ x\cos t - y\operatorname{sen} t\end{pmatrix} = \begin{pmatrix}-\operatorname{sen} t & -\cos t\\ \cos t & \operatorname{sen} t\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

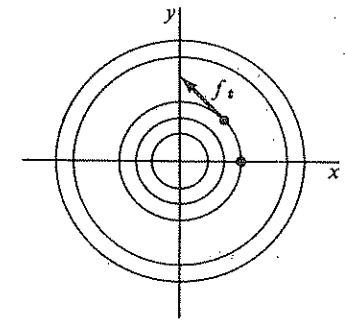


Figura 29

Para calcular $f_t(x, y, t)$ en términos de las coordenadas de posición $(x', y') = f(x, y, t)$, podemos despejar (x, y) de la ecuación

$$\begin{pmatrix}x'\\y'\end{pmatrix} = \begin{pmatrix}\cos t & -\operatorname{sen} t\\ \operatorname{sen} t & \cos t\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

obteniendo

$$\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos t & \operatorname{sen} t\\ \operatorname{sen} t & \cos t\end{pmatrix}\begin{pmatrix}x'\\y'\end{pmatrix}.$$

Por tanto, $f_t(x, y, t)$ está dado por el producto

$$\begin{pmatrix} -\operatorname{sen} t & -\cos t \\ \cos t & \operatorname{sen} t \end{pmatrix}\begin{pmatrix} \cos t & \operatorname{sen} t \\ \operatorname{sen} t & \cos t \end{pmatrix}\begin{pmatrix} x' \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Por tanto, el campo de velocidad v está dado por

$$v(x', y') = \begin{pmatrix} -y' \\ x' \end{pmatrix}.$$

Como v es independiente de t, el flujo es estable.

El Ejemplo 3 ilustra la hipótesis básica que siempre se hace acerca de los flujos: dos partículas no pueden ocupar la misma posición en el mismo instante; en otras palabras, si

$$f(\mathbf{x}_0, t) = f(\mathbf{x}_1, t),$$

entonces $\mathbf{x}_0 = \mathbf{x}_1$. Esta hipótesis no implica que las trayectorias no se cortan (con otras o consigo mismas). Solo se tiene que la ecuación

$$f(\mathbf{x}, t) = \mathbf{y}$$

tiene para t fijo a lo más una solución $\mathbf{x}$ para cada vector dado $\mathbf{y}$. Es claro que una función $\mathbb{R}^4 \xrightarrow{g} \mathbb{R}^3$ de la forma $f(\mathbf{x}, t) = f(x, y, z, t)$ puede ser usada para describir un flujo en tres dimensiones. Los flujos en dimensiones mayores son importantes en la mecánica teórica.

EJERCICIOS

1. Encuentre fórmulas para las derivadas parciales vectoriales $\partial f/\partial x(x, y)$ y $\partial f/\partial y(x, y)$ si

   (a) $f(x, y) = \begin{pmatrix} x + y \\ x - y \\ x^2 + y^2 \end{pmatrix}$.

   (b) $f(x, y) = (e^x \cos y, e^x \operatorname{sen} y)$.

2. Para cada una de las funciones que siguen, encuentre las derivadas parciales vectoriales en los puntos indicados. Dibuje las curvas coordenadas a las cuales estos vectores son tangentes en el punto y dibuje el plano tangente a la imagen de f en el punto.

   (a) $f(x, v) = (u, v, u^2 + v^2)$ para $(1, 1, 2)$.
   (b) $f(u, v) = (\cos u \operatorname{sen} v, \operatorname{sen} u \operatorname{sen} v, \cos v)$ para $(u, v) = (\pi/4, \pi/4)$.

   [*Indicación*: La imagen de f en la parte (b) está en una esfera.]

3. Encuentre una representación paramétrica para el plano tangente del Ejercicio 2(a) y también una representación para la recta perpendicular al plano tangente que pasa por el punto de tangencia.

4. Considere el flujo en dos dimensiones

   $$f(x, y, t) = \begin{pmatrix} x + t \\ y + t^2 \end{pmatrix}, \quad \text{para } t \leq 0.$$

   (a) Dibuje las trayectorias del flujo que empieza en $(x, y) = (0, 0)$, $(0, 1)$ y $(1, 1)$.
   (b) Para $t = 1$ dibuje los vectores velocidad en los puntos $f(x, y, 1)$ con (x, y) elegido como en la parte (a)
   (c) Demuestre que el flujo determinado por f no es estable. [*Indicación*: No es suficiente demostrar que $\partial f/\partial t$ depende de t. Considere $f(0, 0, 1)$ y $\partial f/\partial t(0, 0, 1)$ y compárelos con $f(1, 1, 0)$ y $\partial f/\partial t(1, 1, 0)$.]

5. Considere el flujo en dos dimensiones

   $$f(x, y, t) = (xe^t, ye^t), \quad \text{para } t \geq 0.$$

   (a) Dibuje las trayectorias del flujo que empiezan en $(0, 1)$ y $(1, 1)$.
   (b) Para $t = 1$ dibuje los vectores velocidad en los puntos $f(x, y, 1)$ con (x, y) como en (a).
   (c) De la ecuación $(x', y') = f(x, y, t)$ exprese (x, y) en términos de (x', y') y sustituya el resultado en $\partial f/\partial t(x, y, t)$ para mostrar que el flujo determinado por f es un flujo estable.

6. El flujo

   $$g(x, y, z, t) = \begin{pmatrix} x \cos t - y \operatorname{sen} t \\ x \operatorname{sen} t + y \cos t \\ z + t \end{pmatrix}$$

   es una extensión a tres dimensiones del flujo considerado en el Ejemplo 3. Encuentre $\partial g/\partial t$ y demuestre que el flujo determinado por g es un flujo estable.

7. Límites y continuidad

Los límites y la continuidad han sido considerados, para funciones de una variable, en la Sec. 2 y, para funciones reales, en la Sec. 5. Ahora unificaremos y extenderemos las definiciones y mostraremos cómo construir funciones vectoriales continuas a partir de funciones reales. Para empezar, la definición de límite está basada en la noción de proximidad. Por ejemplo, afirmar que

$$\lim_{x \to 0} \frac{\operatorname{sen} x}{x} = 1$$

es lo mismo que decir que $(\operatorname{sen} x)/x$ está arbitrariamente cerca de 1 cuando x está suficientemente cerca de 0. La proximidad en la recta numérica puede ser expresada mediante desigualdades. Por ejemplo, $|x - 3| < 0{,}4$ significa que la distancia entre el número x y el número 3 es menor que 0,4 o, lo que es lo mismo, que x está en el interior de un intervalo de centro 3 y semilargo 0,4.

La afirmación «(sen x)/x está arbitrariamente cerca de 1 cuando x está suficientemente cerca de 0», se expresa en términos de desigualdades en la forma: Para todo número positivo ϵ, existe un número positivo δ tal que si

$$0 < |x - 0| = |x| < \delta,$$

entonces

$$\left| \frac{\operatorname{sen} x}{x} - 1 \right| < \epsilon.$$

Para la definición de límite $\mathcal{R}^n$, también se necesita disponer de una manera de afirmar que un punto está cercano a otro. Se definirá la distancia usando el largo euclidiano. Para todo $\epsilon > 0$ y todo punto $\mathbf{x}_0$ en $\mathcal{R}^n$, el conjunto de todos los vectores $\mathbf{x}$ de $\mathcal{R}^n$ que satisfacen la desigualdad

$$|\mathbf{x} - \mathbf{x}_0| < \epsilon$$

es una **bola esférica** de radio ϵ y centro $\mathbf{x}_0$. Por ejemplo, si $\mathbf{x}_0 = (1, 2, 1)$, el conjunto de todos los $\mathbf{x}$ en $\mathcal{R}^3$ tales que

$$|\mathbf{x} - \mathbf{x}_0| = \sqrt{(x-1)^2 + (y-2)^2 + (z-1)^2} < 0{,}5$$

es la bola mostrada en la Fig. 30.

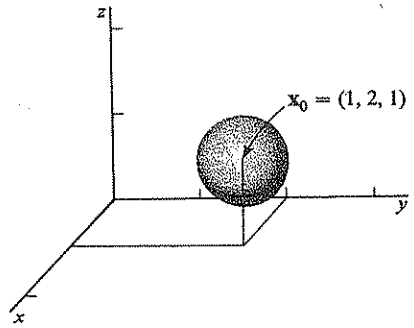


Figura 30

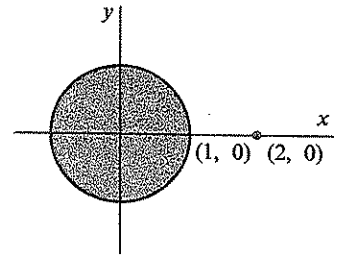


Figura 31

Sea S un subconjunto de $\mathcal{R}^n$ y $\mathbf{x}$ un punto de $\mathcal{R}^n$. Entonces, $\mathbf{x}$ es un punto **límite** de S si para todo $\epsilon > 0$, existe un punto $\mathbf{y}$ en S tal que $0 < |\mathbf{x} - \mathbf{y}| < \epsilon$. Traduciendo al español, esta definición dice que $\mathbf{x}$ es un punto límite de S si en S hay puntos distintos de $\mathbf{x}$ contenidos en una bola de radio arbitrariamente pequeño y de centro en $\mathbf{x}$. Si, por ejemplo, S está formado por el disco definido por $x^2 + y^2 < 1$ y el punto (2, 0), entonces el conjunto de los puntos límites de S está formado por los puntos (x, y) tales que $x^2 + y^2 \leq 1$. Véase Fig. 31. Nótese que el punto (2, 0) no es un punto límite de S, a pesar de que es un punto de S.

Ahora podemos definir el límite de una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$. Sea y_0 un punto

de $\mathcal{R}^m$ y $\mathbf{x}_0$ un punto límite del dominio de f. Un punto arbitrario del dominio de f se denotará por $\mathbf{x}$. Entonces $\mathbf{y}_0$ es el límite de f en $\mathbf{x}_0$ si, para todo $\epsilon > 0$, existe un $\delta > 0$ tal que $|f(\mathbf{x}) - f(\mathbf{y}_0)| < \epsilon$ cuando $\mathbf{x}$ satisface $0 < |\mathbf{x} - \mathbf{x}_0| < \delta$. Esta relación se expresa mediante la ecuación

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0.$$

Un poco menos formalmente, la definición dice que $f(\mathbf{x})$ está arbitrariamente cerca de $\mathbf{y}_0$, siempre que $\mathbf{x}$ esté suficientemente cerca de $\mathbf{x}_0$ y $\mathbf{x} \neq \mathbf{x}_0$. Geométricamente, la idea es la siguiente: Dada una ϵ-bola B_ϵ de centro $\mathbf{y}_0$, existe una δ-bola B_δ de centro $\mathbf{x}_0$ cuya intersección con el dominio de f (con la posible excepción de $\mathbf{x}_0$) es llevado por f en B_ϵ. Un ejemplo en 2 dimensiones está representado en la Fig. 32. La afirmación

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$$

se lee generalmente como «el límite de $f(\mathbf{x})$ cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$ es $\mathbf{y}_0$».

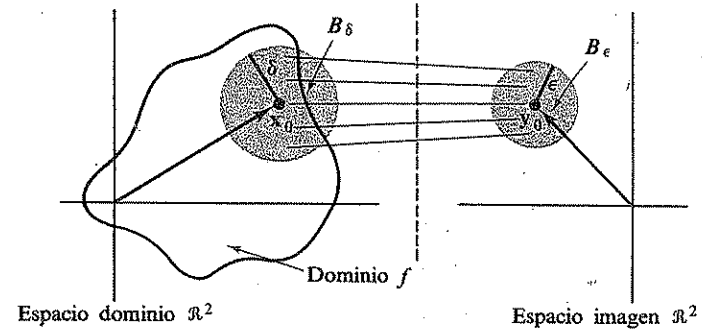


Figura 32

Ejemplo 1. Consideremos la función definida por

$$f(t) = (\cos t, \operatorname{sen} t).$$

El dominio de f es todo $\mathcal{R}$ y, si t_0 es un punto arbitrario de $\mathcal{R}$, f tiene límite $f(t_0)$. Para comprobar esto, usamos propiedades conocidas de cos t y sen t y consideramos

$$\begin{aligned} |f(t) - f(t_0)| &= \sqrt{(\cos t - \cos t_0)^2 + (\operatorname{sen} t - \operatorname{sen} t_0)^2} \\ &\leq |\cos t - \cos t_0| + |\operatorname{sen} t - \operatorname{sen} t_0|. \end{aligned} \tag{1}$$

Esto es válido, ya que $\sqrt{a^2 + b^2} \leq |a| + |b|$. (Véase Ejercicio 15.) Puesto que sen t y cos t son continuas, tenemos que

$$\lim_{t \to t_0} \cos t = \cos t_0, \qquad \text{y} \qquad \lim_{t \to t_0} \operatorname{sen} t = \operatorname{sen} t_0,$$

dado $\epsilon > 0$, podemos elegir $\delta > 0$ tal que

$$|\cos t - \cos t_0| < \frac{\epsilon}{2} \quad \text{y} \quad |\operatorname{sen} t - \operatorname{sen} t_0| < \frac{\epsilon}{2},$$

cuando $|t - t_0| < \delta$. Por tanto, la Ec. (1) muestra que $|f(t) - f(t_0)| < \epsilon$ cuando $|t - t_0| < \delta$.

Ejemplo 2. Consideremos la función real definida en todo $\mathcal{R}^n$ salvo el punto $(x, y) = (0, 0)$, por

$$f(x, y) = \frac{1}{x^2 + y^2}.$$

No existe el límite cuando $(x, y) \to (0, 0)$, por ejemplo, a lo largo de la recta $y = x$. Podemos escribir

$$\lim_{\mathbf{x} \to 0} f(\mathbf{x}) = \infty$$

para describir lo que sucede.

Ejemplo 3. El espacio imagen y el dominio son los mismos que en el ejemplo precedente, y

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}.$$

No existe el límite cuando $(x, y) \to (0, 0)$. Si (x, y) se aproxima a $(0, 0)$, a lo largo de la recta $y = \alpha x$, tenemos

$$\lim_{x \to 0} \frac{x^2 - y^2}{x^2 + y^2} = \lim_{x \to 0} \frac{x^2(1 - \alpha^2)}{x^2(1 + \alpha^2)} = \frac{1 - \alpha^2}{1 + \alpha^2}.$$

Claramente este límite no es independiente de α; él es 0 si $\alpha = 1$ y 1 si $\alpha = 0$.

Las funciones de los dos ejemplos anteriores son todas funciones reales. El teorema siguiente muestra que el problema de existencia y evaluación del límite de una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ se reduce al mismo problema para las funciones coordenadas, las cuales tienen valores reales.

7.1 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ con funciones coordenadas $f_1, \ldots, f_m$ y $\mathbf{y}_0 = (y_1, \ldots, y_n)$ un punto de $\mathcal{R}^n$, entonces

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0 \tag{2}$$

si, y solo si,

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f_i(\mathbf{x}) = y_i, \qquad i = 1, \ldots, m. \tag{3}$$

Demostración. Afirmar que las Ecs. (2) y (3) son equivalentes es lo mismo que afirmar que la distancia

$$|f(\mathbf{x}) - \mathbf{y}_0| = \sqrt{(f_1(\mathbf{x}) - y_1)^2 + \ldots + (f_m(\mathbf{x}) - y_m)^2}$$

puede ser hecha arbitrariamente pequeña si, y solo si, los valores absolutos

$$|f_1(\mathbf{x}) - y_1|, \ldots, |f_m(\mathbf{x}) - y_m|$$

pueden ser hechos arbitrariamente pequeños. Pero la equivalencia de estas dos afirmaciones sigue inmediatamente de las siguientes desigualdades elementales,

$$|f(\mathbf{x}) - \mathbf{y}_0| \geq |f_i(\mathbf{x}) - y_i|, \qquad i = 1, \ldots, m$$

y

$$|f(\mathbf{x}) - \mathbf{y}_0| \leq \sqrt{m} \max_{1 \leq i \leq m} \{|f_i(\mathbf{x}) - y_i|\}.$$

Dejemos la comprobación de estas desigualdades como ejercicio.

Ejemplo 4. Sean f_1 y f_2 las funciones vectoriales definidas por

$$f_1(t) = \begin{pmatrix} t \\ t^2 \\ \operatorname{sen} t \end{pmatrix}, \qquad f_2(t) = \begin{pmatrix} t \\ t^2 \\ \operatorname{sen} \dfrac{1}{t} \end{pmatrix}.$$

Entonces

$$\lim_{t \to 0} f_1(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix};$$

pero lím $f_2(t)$ no existe, ya que la función coordenada sen $(1/t)$ no tiene límite en $t = 0$.

El concepto de continuidad es fundamental en el cálculo. En términos generales, una función es continua si sus valores no cambian abruptamente. En otras palabras, si $\mathbf{x}$ está cerca de $\mathbf{x}_0$, entonces $f(\mathbf{x})$ debe estar cerca de $f(\mathbf{x}_0)$. Esta definición está relacionada con la noción de límite y su definición exacta es la siguiente: **Una función f es continua en $\mathbf{x}_0$** si $\mathbf{x}_0$ está en el dominio de f y $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$. En un punto del dominio de que no es punto límite, también llamado punto **aislado,** la noción de límite no tiene sentido. Sin embargo, por definición, f será continua en tales puntos. Un corolario inmediato del Teorema 7.1 es el siguiente.

7.2 Teorema

Una función vectorial es continua en un punto si, y solo si, sus funciones coordenadas lo son.

Ejemplo 5. La función

$$f_1(t) = \begin{pmatrix} t \\ t^2 \\ \operatorname{sen} t \end{pmatrix}$$

es continua en cualquier valor de t. Por el contrario, la función

$$f_2(t) = \begin{pmatrix} t \\ t^2 \\ \operatorname{sen} \dfrac{1}{t} \end{pmatrix}$$

es continua en todos los puntos, excepto $t = 0$.

Teorema 7.3

Toda función lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^m$ es continua y para tal L existe un número real k tal que

$$|L(\mathbf{x})| \leq k\,|\mathbf{x}|, \quad \text{para todo } \mathbf{x} \text{ en } \mathcal{R}^n.$$

Demostración. Demostraremos primero la desigualdad. Sea $\mathbf{e}_1, \ldots, \mathbf{e}_n$ la base natural de $\mathcal{R}^n$, de modo que, si $\mathbf{x} = (x_1, \ldots, x_n)$, entonces $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. Como L es lineal,

$$L(\mathbf{x}) = x_1L(\mathbf{e}_1) + \ldots + x_nL(\mathbf{e}_n).$$

Por la propiedad de homogeneidad y la desigualdad triangular para la norma, tenemos que

$$|L(\mathbf{x})| \leq |x_1|\,|L(\mathbf{e}_1)| + \ldots + |x_n|\,|L(\mathbf{e}_n)|.$$

Definiendo $k = |L(\mathbf{e}_1)| + \ldots + |L(\mathbf{e}_n)|$ y usando la propiedad $|x_i| \leq |\mathbf{x}|$ para $i = 1, \ldots, n$, se obtiene la desigualdad deseada.

Usamos la desigualdad para demostrar que L es continua. Si $\mathbf{x}$ y $\mathbf{x}_0$ son vectores en $\mathcal{R}^n$, entonces

$$\begin{aligned} |L(\mathbf{x}) - L(\mathbf{x}_0)| &= |L(\mathbf{x} - \mathbf{x}_0)| \\ &\leq k\,|\mathbf{x} - \mathbf{x}_0|. \end{aligned}$$

Esto demuestra que, cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$, $L(\mathbf{x})$ tiende a $L(\mathbf{x}_0)$.

Una función se llama **continua** si ella es continua en todos los puntos de su dominio. Del Teorema 7.2 se concluye que una función $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$ es continua si, y solo si, sus funciones coordenadas $f_1, \ldots, f_n$ son funciones reales continuas de una variable real. Esto incluye a la mayor parte de las funciones del cálculo ordinario, tales como x^2, $\operatorname{sen} x \cdot y$, para $x > 0$, $\log x$. Estas mismas funciones pueden ser usadas para construir ejemplos de funciones coordenadas

que son la base para definir funciones vectoriales de una variable real: $\mathcal{R} \xrightarrow{f} \mathcal{R}^n$. Por ejemplo, las funciones coordenadas de

$$f(x, y) = \left(\frac{\operatorname{sen} xy}{e^{x+y}}, \frac{\cos xy}{e^{x+y}}\right)$$

son continuas. La continuidad de este y otros ejemplos puede ser deducida aplicando repetidamente los tres teoremas siguientes.

7.4 Las funciones $\mathcal{R}^n \xrightarrow{P_k} \mathcal{R}$, donde $P_k(x_1, \ldots, x_n) = x_k$ para $k = 1, \ldots, n$, son continuas.

7.5 Las funciones $\mathcal{R}^2 \xrightarrow{S} \mathcal{R}$ y $\mathcal{R}^2 \xrightarrow{M} \mathcal{R}$ definidas por $S(x, y) = x + y$ y $M(x, y) = xy$, son continuas.

7.6 Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ y $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^p$ son continuas, entonces la composición $g \circ f$ definida por $g \circ f(\mathbf{x}) = g(f(\mathbf{x}))$, es continua en los puntos en que esté definida.

Se deja como ejercicio la demostración de 7.4 y 7.5.

Demostración de 7.6. Sea $\mathbf{x}_0$ un punto límite del dominio de $g \circ f$ y demostraremos que $\lim_{\mathbf{x}\to\mathbf{x}_0} g \circ f(\mathbf{x}) = g \circ f(\mathbf{x}_0)$. Si $\epsilon > 0$, por continuidad de g, podemos encontrar $\delta > 0$ tal que $|g(\mathbf{y}) - g(f(\mathbf{x}_0))| < \epsilon$ si $|\mathbf{y} - f(\mathbf{x}_0)| < \delta$ y $\mathbf{y}$ está en el dominio de g. Como f es continua, podemos encontrar $\delta' > 0$ tal que $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \delta$ si $|\mathbf{x} - \mathbf{x}_0| < \delta$ y $\mathbf{x}$ está en el dominio de f. Por tanto,

$$|g \circ f(\mathbf{x}) - g \circ f(\mathbf{x}_0)| = |g(f(\mathbf{x})) - g(f(\mathbf{x}_0))| < \epsilon$$

para estos mismo vectores $\mathbf{x}y$, por tanto, $g \circ f$ es continua en $\mathbf{x}_0$. Como $\mathbf{x}_0$ era un punto límite arbitrario del dominio de f, el teorema está demostrado.

Ejemplo 6. La función $f(x, y) = \sqrt{1 - x^2 - y^2}$, definida para $|(x, y)| \leq 1$, es continua, ya que puede ser expresada en la forma

$$f(x, y) = \sqrt{1 - (P_1(x, y))^2 - (P_2(x, y))^2},$$

y, por tanto, es la composición de funciones continuas. Análogamente, $g(x, y) = \log(x + y)$, definida para $x + y > 0$, es continua. El producto de f y g, definido por

$$h(x, y) = \sqrt{1 - x^2 - y^2}\,\log(x + y),$$

está definido en el semidisco intersección de los dominios de f y g, mostrado en la Fig. 33. El producto es una función continua, ya que ella es la composición de las funciones vectoriales continuas

$$F(x, y) = (f(x, y), g(x, y))$$

con la función M de 7.5.

Un vector $\mathbf{x}_0$ es un punto interior de un subconjunto de un espacio vectorial si todos los puntos suficientemente cercanos a $\mathbf{x}_0$ también están en el subconjunto. Consideremos, por ejemplo, el subconjunto S de $\mathcal{R}^2$ formado por todos los puntos (x, y) tales que $0 < x \leq 2$ y $-1 \leq y < 1$ [véase Fig. 34(a)]. Los

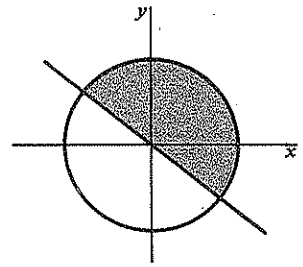


Figura 33

puntos $(1, 0)$, $(\frac{1}{2}, \frac{1}{2})$, $(1, -1)$, $(2, 0)$ pertenecen a S. Los dos primeros son puntos interiores y los dos últimos no. Más generalmente, los puntos interiores de S son exactamente los (x, y) tales que $0 < x < 2$ y $-1 < y < 1$. La definición formal es la siguiente: $\mathbf{x}_0$ es un **punto interior** de un subconjunto S de $\mathcal{R}^n$ si existe un número real positivo δ tal que x pertenece a S si $|\mathbf{x} - \mathbf{x}_0| < \delta$.

Un subconjunto de $\mathcal{R}^n$ es **abierto** si todos sus puntos son interiores. Nótese que, de acuerdo con esta definición, todo el espacio $\mathcal{R}^n$ es abierto. También lo es el subconjunto vacío $\varnothing$ de $\mathcal{R}^n$. Como $\varnothing$ no tiene puntos, la condición de su abierto se satisface en forma vacía. Un conjunto abierto que contiene a un punto dado se llama a menudo una **vecindad** de ese punto.

Ejemplo 7. Para todo $\epsilon > 0$ y todo $\mathbf{x}_0$ en $\mathcal{R}^n$, podemos demostrar que el conjunto B_ϵ de todos los vectores $\mathbf{x}$ tales que $|\mathbf{x} - \mathbf{x}_0| < \epsilon$ es abierto. En

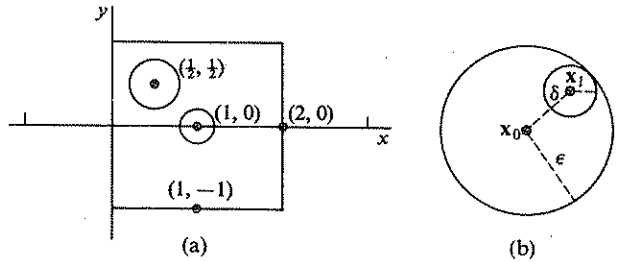


Figura 34

el espacio tridimensional, B_ϵ es la bola de radio ϵ, mostrada en la Fig. 30, para $\epsilon = 0{,}5$. Sea $\mathbf{x}_1$ un punto arbitrario de B_ϵ. Entonces $|\mathbf{x}_1 - \mathbf{x}_0| < \epsilon$. Debemos demostrar que todo vector suficientemente cercano a $\mathbf{x}_1$ está en B_ϵ. Como en la Fig. 34(b), tomemos

$$\delta = \epsilon - |\mathbf{x}_1 - \mathbf{x}_0|.$$

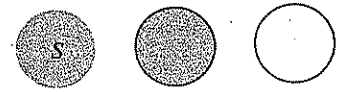
Entonces δ es positivo. Si $\mathbf{x}$ es un vector cualquiera tal que $|\mathbf{x} - \mathbf{x}_1| < \delta$. Por la desigualdad triangular,

$$\begin{aligned} |\mathbf{x} - \mathbf{x}_0| &= |(\mathbf{x} - \mathbf{x}_1) + (\mathbf{x}_1 - \mathbf{x}_0)| \\ &\leq |\mathbf{x} - \mathbf{x}_1| + |\mathbf{x}_1 - \mathbf{x}_0| \\ &< \delta + (\epsilon - \delta) = \epsilon. \end{aligned}$$

Por tanto, $\mathbf{x}$ está en B_ϵ y la demostración queda terminada.

Ejemplo 8. Sea I un conjunto finito de puntos de $\mathcal{R}^n$. Entonces el conjunto de todos los puntos de $\mathcal{R}^n$ que no están en I, es abierto. Por tanto, una función vectorial definida en todo $\mathcal{R}^n$, excepto por un número finito de puntos, tiene como dominio un subconjunto abierto de $\mathcal{R}^n$.

Un conjunto S es **cerrado** si contiene todos sus puntos límites y la **cerradura** de S es el conjunto formado por los puntos de S y los puntos límites de S. La **frontera** de S es el conjunto que se obtiene al quitar a la cerradura de S los puntos interiores de S. Así, el intervalo $a \leq x \leq b$, denotado por $[a, b]$ y llamado intervalo cerrado, es un conjunto cerrado. Un intervalo abierto $a < x < b$, denotado por (a, b) es abierto. El diagrama muestra un conjunto S, su cerradura y su frontera.



EJERCICIOS

En los Ejercicios 1 y 2 tome los dominios de las funciones, lo más grandes que se pueda.

1. ¿En cuáles puntos no tienen límites las funciones siguientes?

(a) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y + \tan x \\ \ln (x + y) \end{pmatrix}$.

(b) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{y}{x^2 + 1} \\ \dfrac{x}{y^2 - 1} \end{pmatrix}$.

(c) $f(x, y) = \dfrac{x}{\operatorname{sen} x} + y$.

(d) $f(x, y) = \begin{cases} \dfrac{x}{\operatorname{sen} x} + y, & \text{si } x \neq 0. \\ 2 + y, & \text{si } x = 0. \end{cases}$

(e) $f(t) = \begin{pmatrix} \operatorname{sen} t \\ \cos t \\ \dfrac{1}{\operatorname{sen} t^2} \end{pmatrix}$.

2. ¿En cuáles puntos no son continuas las siguientes funciones?

(a) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{1}{x^2} + \dfrac{1}{y^2} \\ x^2 + y^2 \end{pmatrix}$.

(b) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 3u - 4v \\ u + 8v \end{pmatrix}$.

(c) $f(x, y) = \begin{cases} \dfrac{x}{\operatorname{sen} x} + y, & \text{si } x \neq 0. \\ 1 + y, & \text{si } x = 0. \end{cases}$

(d) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2}, & \text{si } x^2 + y^2 \neq 0. \\ 0, & \text{si } x^2 + y^2 = 0. \end{cases}$

(e) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \tan u \\ u \sec v \\ v \end{pmatrix}$.

(f) $f(\mathbf{x}) = \dfrac{|\mathbf{x}|}{1 - |\mathbf{x}|^2}$.

3. Si $\mathbf{x}_0 = (1, 2)$, dibuje el conjunto de todos los $\mathbf{x}$ en $\mathcal{R}^2$ tales que

(a) $|\mathbf{x} - \mathbf{x}_0| \leq 3$.
(b) $|\mathbf{x} - \mathbf{x}_0| = 3$.
(c) $|\mathbf{x} - \mathbf{x}_0| < 3$.

4. Indique si es abierto, cerrado, ambos o ninguna de las dos cosas, el subconjunto de todos los vectores $\mathbf{x} = (x, y)$ tales que

(a) $|\mathbf{x} - (1, 2)| \leq 0{,}5$.
(b) $|\mathbf{x} - (1, 2)| < 0{,}5$.
(c) $|\mathbf{x} - (1, 2)| < -0{,}5$.
(d) $0 < x < 3$ y $0 < y < 2$.
(e) $2 \leq x < 3$ y $0 < y < 2$.
(f) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} < 1$.

(g) $\mathbf{x} \neq (0, 2)$ o $(1, 2)$.
(h) $x^2 + y^2 > 0$.
(i) $x > 0$.
(j) $x > y$.

5. Sea S el conjunto de todos los puntos (x, y) en $\mathcal{R}^2$ que satisfacen $0 < x^2 + y^2 < 1$ y de todos los puntos del eje x que satisfacen $1 \leq x < 2$.
(a) Describa la frontera de S.
(b) ¿Cuáles son los puntos interiores de S?
(c) Describa la cerradura de S.

6. (a) Demuestre que la unión de una colección arbitraria de subconjuntos abiertos de $\mathcal{R}^n$ es abierta.
(b) Demuestre que la intersección de una familia *finita* de subconjuntos abiertos de $\mathcal{R}^n$ es abierta.
(c) Dé un ejemplo de que la intersección no vacía de *infinitos* subconjuntos abiertos de $\mathcal{R}^n$ puede no ser abierta.

7. Demuestre que $\mathbf{x}$ es un punto de la frontera de un conjunto S si, y solo si, toda vecindad de $\mathbf{x}$ contiene un punto de S y un punto del complemento de S.

8. Se dice que una función f tiene una **discontinuidad removible** si (a) f no es continua en $\mathbf{x}_0$, (b) existe un vector $\mathbf{y}_0$ tal que $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$. Dé ejemplos de funciones f y un punto $\mathbf{x}_0$ tales que f no es continua en $\mathbf{x}_0$ y (1) f tiene una discontinuidad removible de $\mathbf{x}_0$, (2) f no tiene una discontinuidad removible en $\mathbf{x}_0$.

9. Demuestre que toda traslación es una función vectorial continua. Una función vectorial $\mathcal{R}^n \xrightarrow{t} \mathcal{R}^n$ es una **traslación** si existe un vector $\mathbf{y}_0$ en $\mathcal{R}^n$ tal que $t(\mathbf{x}) = \mathbf{x} + \mathbf{y}_0$ para todo $\mathbf{x}$ en $\mathcal{R}^n$.

10. Demuestre 7.4 y 7.5.

11. Si f y g son funciones vectoriales con el mismo dominio y la misma imagen, demuestre que

$$\lim_{\mathbf{x} \to \mathbf{x}_0} (f(\mathbf{x}) + g(\mathbf{x})) = \lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) + \lim_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x}),$$

siempre que $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x})$ y $\lim_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x})$ existan.

12. Sea L una recta y P un plano en $\mathcal{R}^3$. ¿Es L o P un subconjunto abierto?

13. Sea S un subconjunto cerrado de $\mathcal{R}^n$. Demuestre que el complemento de S en $\mathcal{R}^n$ es abierto.

14. Converso del Ejercicio 13. Si S es un subconjunto abierto de $\mathcal{R}^n$, demuestre que el complemento de S en $\mathcal{R}^n$ es abierto.

15. Demuestre que $\sqrt{a^2 + b^2} \leq |a| + |b|$.

16. Demuestre las desigualdades

$$|(\mathbf{z}_1, \ldots, \mathbf{z}_m)| \geq |\mathbf{z}_i|, \qquad i = 1, \ldots, m$$

y

$$|(\mathbf{z}_1, \ldots, \mathbf{z}_m)| \leq \sqrt{m} \max \{|\mathbf{z}_i|\}.$$

17. Demuestre que el Teorema 7.3 también puede ser demostrado usando los Teoremas 7.4 y 7.5

8. Diferenciales

Muchas de las técnicas del cálculo están basadas en la idea de aproximar una función vectorial por una función lineal o por una función afín. Recordemos que una función $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$ es **afín** si existe una función lineal $\mathbb{R}^n \xrightarrow{L} \mathbb{R}^m$ y un vector $\mathbf{y}_0$ en $\mathbb{R}^m$ tal que $A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0$ para todo $\mathbf{x}$ en $\mathbb{R}^n$.

Veremos que las funciones afines son la base del cálculo diferencial de las funciones vectoriales.

Ejemplo 1. Una función afín está definida por ecuaciones lineales en las coordenadas de las variables. Por ejemplo, consideremos el punto

$$\mathbf{y}_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

y la función lineal $\mathbb{R}^3 \xrightarrow{L} \mathbb{R}^2$, definida por la matriz

$$\begin{pmatrix} 2 & 3 & 0 \\ 1 & 0 & 5 \end{pmatrix}.$$

La función afín $A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0$, donde

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

está descrita por las ecuaciones

$$u = 2x + 3y + 1$$
$$v = x + 5z + 2$$
$$w = x + 2y.$$

Como todo sistema de ecuaciones lineales puede ser resuelto en forma sistemática, toda función afín puede ser analizada de la misma manera.

Estudiaremos la posibilidad de aproximar una función vectorial f arbitraria cerca de un punto $\mathbf{x}_0$ de su dominio, por una función $\mathbf{x}_0$, una función que puede ser muy complicada por una muy simple. Antes de decidir acerca de la posibilidad de la existencia de una aproximación, debemos precisar lo que entendemos por aproximación. Empezamos exigiendo que $f(\mathbf{x}_0) = A(\mathbf{x}_0)$. Como $A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0$, donde L es lineal, tenemos que $f(\mathbf{x}_0) = L(\mathbf{x}_0) + \mathbf{y}_0$ y, por tanto,

$$A(\mathbf{x}) = L(\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x}_0). \tag{1}$$

Un requisito que parece muy natural es

$$\lim_{\mathbf{x}\to\mathbf{x}_0} (f(\mathbf{x}) - A(\mathbf{x})) = 0. \tag{2}$$

Sin embargo, la Ec. (2) parece decir más de lo que realmente dice. Para ver lo que dice realmente, notemos que de (1) se tiene que

$$f(\mathbf{x}) - A(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0).$$

Ahora bien, toda función lineal es continua, de modo que $\lim_{\mathbf{x}\to\mathbf{x}_0} L(\mathbf{x} - \mathbf{x}_0) = L(0) = 0$. Por tanto,

$$\lim_{\mathbf{x}\to\mathbf{x}_0} (f(\mathbf{x}) - A(\mathbf{x})) = \lim_{\mathbf{x}\to\mathbf{x}_0} (f(\mathbf{x}) - f(\mathbf{x}_0)).$$

De aquí se deduce que la Ec. (2) dice exactamente que la función vectorial f es continua en $\mathbf{x}_0$. Esto es significativo, pero no dice nada acerca de L. Por tanto, para que nuestra noción de aproximación sea capaz de distinguir una función afín de otra o mida de alguna manera lo bien que A aproxima a f, se necesita algún requisito adicional. Una condición natural y que nosotros exigiremos es que $f(\mathbf{x}) - A(\mathbf{x})$ tienda a cero más rápidamente que lo que $\mathbf{x}$ tiende a $\mathbf{x}_0$. Es decir, exigimos que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|} = 0.$$

o, lo que es lo mismo, podemos pedir que f pueda ser representada en la forma

$$f(\mathbf{x}) = f(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + |\mathbf{x} - \mathbf{x}_0| Z(\mathbf{x} - \mathbf{x}_0),$$

donde $Z(\mathbf{y})$ es una función que tiende a cero cuando $\mathbf{y}$ tiende a cero.

Una función $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$ se dice **diferenciable** en $\mathbf{x}_0$ si

1. $\mathbf{x}_0$ es un punto interior del dominio de f.
2. Existe una función afín que aproxima a f cerca de $\mathbf{x}_0$. Es decir, existe una función lineal $\mathbb{R}^n \xrightarrow{L} \mathbb{R}^m$ tal que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|} = 0.$$

La función lineal L se llama la **diferencial** de f en $\mathbf{x}_0$. Si f es diferenciable en cada punto de su dominio, se dice que ella es **diferenciable.**

De acuerdo con nuestra definición, el dominio de una función diferenciable es un conjunto abierto. Sin embargo, es conveniente extender la noción de diferenciabilidad a funciones f definidas en un subconjunto arbitrario S del espacio dominio. Hacemos esto diciendo que f es diferenciable si ella es la restricción a S de una función diferenciable cuyo dominio es un conjunto abierto.

Ejemplo 2. La función f definida por $f(x, y) = \sqrt{1 - x^2 - y^2}$ tiene como dominio el disco $x^2 + y^2 \leq 1$. Su gráfico se muestra en la Fig. 35. Los puntos interiores de su dominio son los pares (x, y) tales que $x^2 + y^2 < 1$. Veremos que esta función es diferenciable en todos esos puntos.

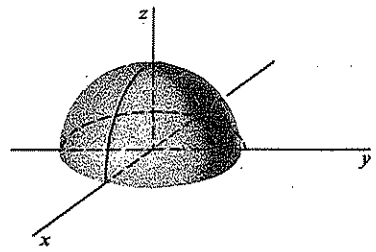


Figura 35

En dimensión 1 una función afín tiene la forma $ax + b$. Por tanto, una función real de la variable x es diferenciable en x_0 si puede ser aproximada cerca de x_0 por la función $A(x) = ax + b$. Como $f(x_0) = A(x_0) = ax_0 + b$, tenemos que

$$A(x) = ax + b = a(x - x_0) + f(x_0).$$

La parte lineal de A (denotada anteriormente por L) es, en este caso, la multiplicación por el número real a. La norma euclideana de un número real es su valor absoluto, de modo que la condición (2) de la definición de diferenciabilidad se transforma en

$$\lim_{x \to x_0} \frac{f(x_j) - f(x_0) - a(x - x_0)}{|x - x_0|} = 0.$$

Esto es equivalente a

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = a.$$

El número a se denota generalmente por $f'(x_0)$ y se llama la **derivada de** f en x_0. La función afín A está dada por

$$A(x) = f(x_0) + f'(x_0)(x - x_0).$$

Su gráfico es la **recta tangente** al gráfico de f en x_0. La Fig. 36 muestra un ejemplo típico. Por tanto, hemos visto que la definición general de diferenciabilidad de funciones vectoriales se reduce, en el caso de dimensión 1, a la definición habitual de los cursos de cálculo en una variable.

Una función lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^m$ puede ser representada por una matriz n-por-m. Veremos más adelante que la matriz de cualquier función lineal L que satisface las condiciones 1 y 2 de la definición de diferenciabilidad puede ser calculada en términos de las derivadas parciales de f. Por tanto, L queda determinada unívocamente por f en cada punto interior de su dominio. Por tanto, podemos hablar de *la* diferencial de f en $\mathbf{x}_0$ y la denotaremos por $d_{\mathbf{x}_0}f$.

Para encontrar la matriz de $d_{\mathbf{x}_0}f$, donde $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$, consideremos la base

natural $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ del espacio dominio $\mathcal{R}^n$. Si $\mathbf{x}_0$ es un punto interior del dominio de f, los vectores

$$\mathbf{x}_j = \mathbf{x}_0 + t\mathbf{e}_j, \qquad j = 1, \ldots, n,$$

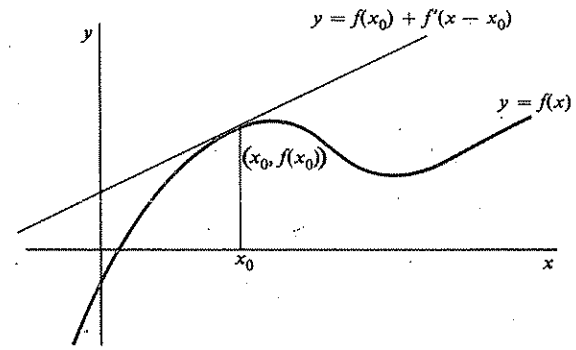


Figura 36

están en el dominio de f para t suficientemente pequeño. (¿Por qué?) Por la condición (2) de la definición de la diferenciabilidad de f, tenemos que

$$\lim_{t \to 0} \frac{f(\mathbf{x}_j) - f(\mathbf{x}_0) - d_{\mathbf{x}_0}f(t\mathbf{e}_j)}{t} = 0,$$

para $j = 1, \ldots, n$. Como $d_{\mathbf{x}_0}f$ es una función lineal, se tiene que

$$\lim_{t \to 0} \frac{f(\mathbf{x}_j) - f(\mathbf{x}_0)}{t} = d_{\mathbf{x}_0}f(\mathbf{e}_j),$$

para $j = 1, \ldots, n$. Pero, por el Teorema 4.2 del Cap. 1, $d_{\mathbf{x}_0}f(\mathbf{e}_j)$ es la columna j de la matriz de $d_{\mathbf{x}_0}f$. Por otra parte, el vector $\mathbf{x}_j$ es distinto de $\mathbf{x}_0$ solo en la j-ésima coordenada y la diferencia en esta coordenada es el número t. Por tanto, el término de la izquierda de la última ecuación es precisamente la derivada parcial $(\partial f/\partial x_j)(\mathbf{x}_0)$, y resulta ser, entonces, la columna j de la matriz de $d_{\mathbf{x}_0}f$. Si las funciones coordenadas de f son $f_1, \ldots, f_m$, entonces

$$\frac{\partial f}{\partial x}(\mathbf{x}_0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(\mathbf{x}_0) \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial f_m}{\partial x_j}(\mathbf{x}_0) \end{pmatrix},$$

y la matriz completa de $d_{\mathbf{x}_0}f$ es

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

Esta matriz se llama la **matriz jacobiana** o **derivada** de f en $\mathbf{x}_0$ y se denota por $f'(\mathbf{x}_0)$. Podemos resumir lo hecho como sigue.

8.1 Teorema

Si la función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es diferenciable en $\mathbf{x}_0$, entonces su diferencial $d_{\mathbf{x}_0}f$ está determinada unívocamente y su matriz es la matriz jacobiana de f. Es decir, para todo $\mathbf{y}$ en $\mathcal{R}^n$

$$d_{\mathbf{x}_0}f(\mathbf{y}) = f'(\mathbf{x}_0)\mathbf{y}.$$

A pesar de que la transformación lineal $d_{\mathbf{x}_0}f$ y su matriz $f'(\mathbf{x}_0)$ son lógicamente distintas, la ecuación anterior muestra que ellas pueden ser identificadas en la práctica, siempre que se subentienda que la matriz de $d_{\mathbf{x}_0}f$ se toma con respecto de las bases naturales de $\mathcal{R}^n$ y $\mathcal{R}^m$.

Ejemplo 3. La función

$$f(x, y, z) = \begin{pmatrix} x^2 + e^y \\ x + y \operatorname{sen} z \end{pmatrix}$$

tiene funciones coordenadas $f_1(x, y, z) = x^2 + e^y$ y $f_2(x, y, z) = x + y \operatorname{sen} z$. La matriz jacobiana en (x, y, z) es, por definición,

$$f'(x, y, z) = \left(\frac{\partial f}{\partial x}(x, y, z)\, \frac{\partial f}{\partial y}(x, y, z)\, \frac{\partial f}{\partial z}(x, y, z)\right),$$

de modo que, en este ejemplo,

$$f'(x, y, z) = \begin{pmatrix} 2x & e^y & 0 \\ 1 & \operatorname{sen} z & y \cos z \end{pmatrix}.$$

Por tanto, la diferencial de f en $(1, 1, \pi)$ es la función lineal cuya matriz es

$$f'(1, 1, \pi) = \begin{pmatrix} 2 & e & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Ejemplo 4. La función f definida por

$$f(x, y) = (x^2 + 2xy + y^2, xy^2 + x^2y)$$

tiene diferencial $d_{\mathbf{x}}f$ en $\mathbf{x} = (x, y)$ representada por la matriz jacobiana

$$f'(x, y) = \begin{pmatrix} 2x + 2y & 2x + 2y \\ y^2 + 2xy & 2xy + x^2 \end{pmatrix}.$$

¿Cómo se puede saber si una función vectorial es diferenciable? El Teorema 8.1 dice solo que si f es diferenciable, entonces su diferencial está representado por la matriz jacobiana. No va en la dirección opuesta. Por tanto, en los Ejemplos 3 y 4 no se puede decidir si las funciones consideradas son diferenciables. Así como la derivada de una función real puede no existir, del mismo modo en general, una función vectorial puede no ser diferenciable. El teorema que sigue es un criterio conveniente para la diferenciabilidad.

8.2 Teorema

Si el dominio de $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es un conjunto abierto D en el cual todas las derivadas parciales $\partial f_i/\partial x_j$ de las funciones coordenadas de f son continuas, entonces f es diferenciable en todo punto de D.

Demostración. Sea L la función lineal definida por la matriz jacobiana de f. El teorema habrá sido demostrado si se puede demostrar que L satisface

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|} = 0. \tag{3}$$

Como, por el Teorema 7.1, una función vectorial tiene un límite si, y solo si, las funciones coordenadas tienden a las coordenadas del límite, es suficiente demostrar el teorema para las funciones coordenadas de f o, lo que es lo mismo, demostrarlo cuando f es una función real. Si

$$\mathbf{x} = (x_1, \ldots, x_n)$$

y

$$\mathbf{x}_0 = (a_1, \ldots, a_n),$$

hagamos

$$\mathbf{y}_k = (x_1, \ldots, x_k, a_{k+1}, \ldots, a_n), \qquad k = 0, 1, \ldots, n,$$

de modo que $\mathbf{y}_0 = \mathbf{x}_0$ y $\mathbf{y}_n = \mathbf{x}$. La Fig. 37 muestra estos vectores en tres di-

mensiones. Entonces, como en la suma que sigue los términos sucesivos se cancelan, solo quedan el primer y último término y tenemos que

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \sum_{k=1}^{n} (f(\mathbf{y}_k) - f(\mathbf{y}_{k-1})).$$

Como $\mathbf{y}_k$ y $\mathbf{y}_{k-1}$ solo difieren en sus k-ésimas coordenadas, podemos usar el teorema del valor medio para funciones reales de una variable y obtenemos

$$f(\mathbf{y}_k) - f(\mathbf{y}_{k-1}) = (x_k - a_k) \frac{\partial f}{\partial x_k}(\mathbf{z}_k),$$

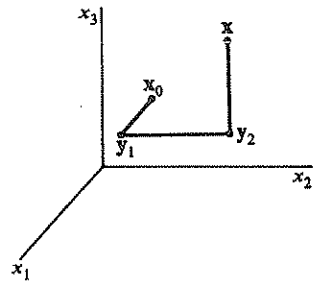


Figura 37

donde $\mathbf{z}_k$ es un punto en el segmento de recta que une a $\mathbf{y}_k$ con $\mathbf{y}_{k-1}$. Entonces

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \sum_{k=1}^{n} (x_k - a_k) \frac{\partial f}{\partial x_k}(\mathbf{z}_k).$$

También tenemos que, por definición,

$$L(\mathbf{x} - \mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \cdots \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) \begin{pmatrix} x_1 - a_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n - a_n \end{pmatrix}$$

$$= \sum_{k=1}^{n} (x_k - a_k) \frac{\partial f}{\partial x_k}(\mathbf{x}_0).$$

Por tanto,

$$|f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)| = \left| \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(\mathbf{z}_k) - \frac{\partial f}{\partial x_k}(\mathbf{x}_0) \right)(x_k - a_k) \right|$$

$$\leq \sum_{k=1}^{n} \left| \frac{\partial f}{\partial x_k}(\mathbf{z}_k) - \frac{\partial f}{\partial x_k}(\mathbf{x}_0) \right| |\mathbf{x} - \mathbf{x}_0|,$$

donde hemos usado la desigualdad triangular y el hecho que

$$|x_k - a_k| \leq |\mathbf{x} - \mathbf{x}_0| \quad \text{para} \quad k = 1, 2, \ldots, n.$$

Como hemos supuesto que las derivadas parciales son continuas en x_0 y $\mathbf{z}_k$ tiende a $\mathbf{x}_0$ cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$, se tiene el límite de la Ec. (4).

Los elementos de las matrices jacobianas que se obtuvieron en los Ejemplos 3 y 4 son funciones continuas. Como consecuencia del teorema, las funciones de esos ejemplos son diferenciables.

Ejemplo 5. La función

$$f(x, y) = \sqrt{1 - x^2 - y^2}$$

definida para todo (x, y) tal que $x^2 + y^2 < 1$ es la del Ejemplo 2, excepto que hemos eliminado la frontera del disco, de modo que su dominio es un conjunto abierto. La matriz jacobiana es

$$(f_x \quad f_y) = \left(\frac{-x}{\sqrt{1 - x^2 - y^2}} \quad \frac{-y}{\sqrt{1 - x^2 - y^2}} \right).$$

Los elementos son continuos en el disco abierto, de modo que, por el Teorema 8.2, f es diferenciable allí.

Ejemplo 6. Consideremos la función

$$f(t) = \begin{pmatrix} \cos t \\ \operatorname{sen} t \end{pmatrix}, \qquad -\infty < t < \infty.$$

La derivada $f'(t_0)$ es la matriz 2-por-1

$$\begin{pmatrix} -\operatorname{sen} t_0 \\ \cos t_0 \end{pmatrix}.$$

Es instructivo considerar la matriz como un vector en el espacio imagen de f y dibujarlo con su cola en el punto $f(t_0)$. Para $t_0 = 0$, $\pi/4$, $\pi/3$, $\pi/2$ y π las matrices respectivas de la diferencial $d_{t_0}f$ son

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -\sqrt{2/2} \\ \sqrt{2/2} \end{pmatrix}, \quad \begin{pmatrix} -\sqrt{3/2} \\ 1/2 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \text{y} \quad \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

La Fig. 38 muestra estos vectores con sus colas en los puntos correspondientes. Evidentemente, por lo menos para curvas, la diferencial está relacionada con la noción de un vector tangente. La función afín que mejor aproxima a $f(t)$ en una vecindad de t_0 es la función vectorial de t definida por

$$f'(t_0)(t - t_0) + f(t_0),$$

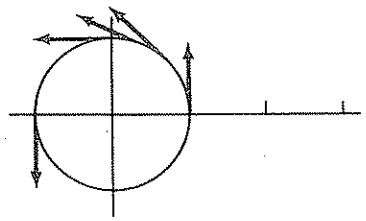


Figura 38

que, en términos de matrices, se escribe en la forma

$$t\begin{pmatrix}-\operatorname{sen} t_0\\ \cos t_0\end{pmatrix}+\begin{pmatrix}\cos t_0 + t_0 \operatorname{sen} t_0\\ \operatorname{sen} t_0 - t_0 \cos t_0\end{pmatrix}.$$

Esta es la ecuación de una recta tangente a la imagen de f en $f(t_0)$.

Una buena interpretación geométrica de la diferencial de una función real $f(x, y)$ en $\mathbf{x}_0 = (x_0, y_0)$ se obtiene mirando a la superficie definida explícitamente por f y al plano tangente a esta superficie en $(x_0, y_0 f(\mathbf{x}_0))$. La Fig. 39

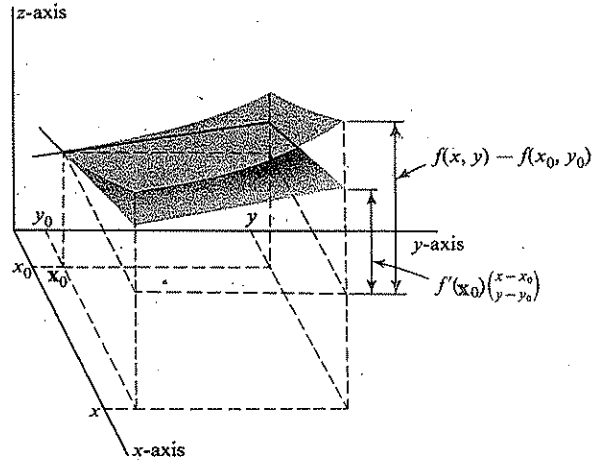


Figura 39

muestra un ejemplo. El plano tangente es el gráfico de la función afín definida por

$$A(x, y) = f(x_0, y_0) + f'(x_0, y_0)\begin{pmatrix}x - x_0\\ y - y_0\end{pmatrix}, \quad \text{todo } (x, y) \text{ en } \mathbb{R}^2.$$

La diferencia $A(x, y) - f(x_0, y_0)$ es una buena aproximación del incremento $f(x, y) - f(x_0, y_0)$, siempre que (x, y) esté cerca de (x_0, y_0). La Fig. 39 es análoga a las figuras incluidas en muchos libros de cálculo de una variable para ilustrar la diferencial de una función real de una variable.

El ejemplo que daremos en el Ejercicio 19 muestra que la continuidad de la derivada f' no es necesaria para la diferenciabilidad de f. Sin embargo, la hipótesis del Teorema 8.2 es lo suficientemente importante como para darle un nombre. Una función vectorial f es **continuamente diferenciable** en un conjunto abierto D si los elementos de la matriz jacobiana de f son continuas en D. Por tanto, las funciones de los Ejemplos 2-6 no solo son diferenciables, sino que también continuamente diferenciables en el interior de sus dominios.

Finalmente, hacemos notar que *si una función f es diferenciable, entonces ella es continua* (a pesar de que los elementos de f' pueden no serlo). La demostración es muy sencilla y la dejaremos como un ejercicio. (Véase Ejercicio 11.)

EJERCICIOS

1. Una función lineal $\mathbb{R}^n \xrightarrow{L} \mathbb{R}^m$ está definida por una matriz (a_{ij}) y

$$\mathbf{y}_0 = \begin{pmatrix}b_1\\ \cdot\\ \cdot\\ \cdot\\ b_m\end{pmatrix}$$

es un vector en $\mathbb{R}^m$. Construya un ejemplo eligiendo m y n, una matriz (a_{ij}) y un vector

$$\begin{pmatrix}b_1\\ \cdot\\ \cdot\\ \cdot\\ b_m\end{pmatrix}$$

de modo que la función afín

$$A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0, \qquad \text{todo } \mathbf{x} \text{ en } \mathbb{R}^n,$$

(a) Define explícitamente una recta en $\mathbb{R}^2$.
(b) Define explícitamente una recta en $\mathbb{R}^3$.
(c) Define explícitamente un plano en $\mathbb{R}^3$.
(d) Define paramétricamente una recta en $\mathbb{R}^3$.
(e) Define paramétricamente un plano en $\mathbb{R}^3$.

¿Qué condición debe satisfacer la matriz (a_{ij}) para dar un ejemplo como en (e)?

2. Sea f la función vectorial definida por

$$f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x^2 - y^2\\ 2xy\end{pmatrix},$$

Encuentre la derivada de f en los siguientes puntos:

(a) $\begin{pmatrix}x\\y\end{pmatrix}$, (b) $\begin{pmatrix}a\\b\end{pmatrix}$, (c) $\begin{pmatrix}1\\0\end{pmatrix}$, (d) $\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$.

$$\left[\textit{Resp}: \text{(c)} \quad \begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}.\right]$$

3. Encuentre la derivada de cada una de las funciones siguientes en los puntos indicados.

(a) $f\begin{pmatrix}x\\y\end{pmatrix} = x^2 + y^2$ a $\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$.

(b) $g(x, y, z) = xyz$ a $(x, y, z) = (1, 0, 0)$.

(c) $f(t) = \begin{pmatrix}\operatorname{sen} t\\ \cos t\end{pmatrix}$ a $= \dfrac{\pi}{4}$ $\left[\textit{Resp}: \begin{pmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{pmatrix}.\right]$

(d) $f(t) = \begin{pmatrix}e^t\\ t\\ t^2\end{pmatrix}$ a $= 1$.

(e) $g(x, y) = \begin{pmatrix}x + y\\ x^2 + y^2\end{pmatrix}$ a $(x, y) = (1, 2)$.

(f) $A\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}u + v\\ u - v\\ 1\end{pmatrix}$ a $\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$.

(g) $T\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}u\cos v\\ u \operatorname{sen} v\\ v\end{pmatrix}$ a $\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}1\\\pi\end{pmatrix}$. $\left[\textit{Resp}: \begin{pmatrix}-1 & 0\\ 0 & -1\\ 0 & 1\end{pmatrix}.\right]$

(h) $f(x, y, z) = (x + y + z, xy + yz + zx, xyz)$ a (x, y, z).

4. Sea P una transformación del espacio euclidiano 3-dimensional en el espacio 2-dimensional, definida por $P(x, y, z) = (x, y)$.

(a) ¿Cuál es la interpretación geométrica de esta transformación?

(b) Demuestre que P es diferenciable en todos los puntos y encuentre la matriz de la diferencial de P en $(1, 1, 1)$. $\left[\textit{Resp}: \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}.\right]$

5. (a) Dibuje la curva en $\mathcal{R}^2$ definida paramétricamente por la función

$$g(t) = (t - 1, t^2 - 3t + 2), \qquad -\infty < t < \infty.$$

(b) Encuentre la función afín que aproxima a g (1) cerca de $t = 0$, (2) cerca de $t = 2$. [*Resp.* $A(t) = (t - 1, t - 2)$.]

(c) Dibuje la curva definida paramétricamente por la función afín.

6. Sea f la función dada en el Ejercicio 2 y sean

$$\mathbf{x}_0 = \begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{y}_1 = \begin{pmatrix}0,1\\0\end{pmatrix}, \quad \mathbf{y}_2 = \begin{pmatrix}0\\1,0\end{pmatrix}, \quad \mathbf{y}_3 = \begin{pmatrix}0,1\\0,1\end{pmatrix}.$$

(a) Calcule $f(\mathbf{x}_0 + \mathbf{y}_i)$ para $i = 1, 2, 3$.

(b) Encuentre la función afín A que aproxima a f cerca de $\mathbf{x}_0$.

(c) Use A para encontrar aproximaciones de los vectores

$$f(\mathbf{x}_0 + \mathbf{y}_i), \qquad i = 1, 2, 3.$$

7. (a) Dibuje la superficie en $\mathcal{R}^3$ definida explícitamente por la función

$$f(x, y) = 4 - x^2 - y^2.$$

(b) Encuentre la función afín que aproxima a f (1) cerca de $(0, 0)$, (2) cerca de $(2, 0)$. [*Resp.* $A(x, y) = 8 - 4x$.]

(c) Dibuje los gráficos de las funciones afines de (b).

8. ¿Cuál es la derivada de la función afín

$$\begin{pmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} + \begin{pmatrix}a_0\\b_0\\c_0\end{pmatrix}?$$

9. Demuestre que toda función lineal es su propia diferencial.

10. Demuestre que si la función f es diferenciable en $\mathbf{x}_0$, entonces

$$f'(\mathbf{x}_0)\mathbf{x} = \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{x}) - f(\mathbf{x}_0)}{t}.$$

11. Demuestre que si una función vectorial es diferenciable en un punto, ella es continua en ese punto. [*Indicación*: Multiplique el cociente de la definición de la diferencial por $|\mathbf{x} - \mathbf{x}_0|$.]

12. ¿En cuáles puntos no son diferenciables las siguientes funciones? ¿Por qué?

(a) $f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\dfrac{1}{x^2} + \dfrac{1}{y^2}\\ x^2 + y^2\end{pmatrix}$.

(b) $g\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\sqrt{x^2 - y^2}\\ x + y\end{pmatrix}$.

(c) $f(u, v) = |u + v|$.

(d) $h(x, y) = \begin{cases} (x \operatorname{sen}(1/x),\ x^2 + y^2) & \text{si } x \neq 0. \\ (0,\ x^2 + y^2), & \text{si } x = 0. \end{cases}$

13. Demuestre que toda traslación es diferenciable. ¿Cuál es su diferencial?

14. Consideremos la función $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ definida por $f(\mathbf{x}) = |\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$. Demuestre que $f'(\mathbf{x})\mathbf{y} = 2\mathbf{x} \cdot \mathbf{y}$ para todo $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$.

15. ¿Es diferenciable en todos los puntos de su dominio la función $\mathbb{R}^n \xrightarrow{g} \mathbb{R}$ definida por $g(\mathbf{x}) = |\mathbf{x}|$?

16. Consideremos la función

$$N(\mathbf{x}) = \max\{|x_1|, \ldots, |x_n|\},$$

para $\mathbf{x} = (x_1, \ldots, x_n)$ en $\mathbb{R}^n$. ¿En cuáles puntos no es diferenciable esta función? Dé una respuesta para

(a) $n = 1$,
(b) $n = 2$,
(c) n arbitrario.

17. Demuestre que si f y g son diferenciables en $\mathbf{x}_0$ y a es un número real, entonces $f + g$ y af son diferenciables en $\mathbf{x}_0$ y

(a) $d_{\mathbf{x}_0}(f + g) = d_{\mathbf{x}_0}f + d_{\mathbf{x}_0}g$.
(b) $d_{\mathbf{x}_0}(af) = a(d_{\mathbf{x}_0}f)$.

El dominio de $f + g$ es la intersección de los dominios de f y de g. [*Indicación*: Use la unicidad de la diferencial.]

18. Verifique que la función

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 - y^2}, & x \neq \pm y, \\ 0, & x = \pm y, \end{cases}$$

tiene una matriz jacobiana en $(0, 0)$, pero no es diferenciable en este punto.

19. Demuestre que la función definida por

$$f(x) = \begin{cases} x^2 \operatorname{sen} \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

es diferenciable para todo x, pero que no es continuamente diferenciable en $x = 0$.

9. El método de Newton

En esta sección consideraremos el método de Newton para encontrar una solución aproximada de la ecuación $f(\mathbf{x}) = 0$, donde $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n$ es una función no lineal. El caso lineal ha sido discutido en el Cap. 2. Empezamos considerando la idea de aproximar un vector en $\mathbb{R}^n$ por los términos de una sucesión de vectores en $\mathbb{R}^n$. Estamos acostumbrados a considerar un número real, como, por ejemplo, $\sqrt{2}$, como aproximado por una sucesión de números racionales, tal como 1; 1,4; 1,41; 1,414; ... Esta idea se extiende inmediatamente al caso de vectores.

Ejemplo 1. Consideremos el vector $(\sqrt{2}, \pi)$ en $\mathbb{R}^2$. Supongamos que $\sqrt{2}$ es aproximado por una sucesión 1; 1,4; 1,41; 1,414; ... y que π es aproximado por 3; 3,1; 3,14; 3,141; ... Entonces podemos considerar la sucesión de vectores (1,3); (1,4, 3,1); (1,41, 3,14); (1,414, 3,141); ... como una aproximación del vector $(\sqrt{2}, \pi)$. Es natural aparejar los números en la forma que lo hemos hecho y no en otra, ya que las coordenadas de cada par son aproximaciones de $\sqrt{2}$ y π con el número de decimales dados.

Para precisar estas ideas definimos el límite de una sucesión en $\mathbb{R}^n$. Sea $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, ... una sucesión infinita de vectores en $\mathbb{R}^n$ en que a cada entero positivo se le asocia un vector único. Supongamos que existe un vector $\mathbf{x}$ en $\mathbb{R}^n$ tal que para todo $\epsilon > 0$ existe un entero N para el cual

$$|\mathbf{x}_k - \mathbf{x}| < \epsilon$$

siempre que $k \geq N$. Entonces decimos que la sucesión **converge** al **límite** $\mathbf{x}$ y escribimos

$$\lim_{k \to \infty} \mathbf{x}_k = \mathbf{x}.$$

Podemos resumir esto diciendo que la sucesión $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ converge a $\mathbf{x}$ si $|\mathbf{x}_k - \mathbf{x}|$ es arbitrariamente pequeño para todo k suficientemente grande. La Fig. 40 muestra una sucesión cuyos términos están a una distancia menor que ϵ si $k \geq 6$. Dejamos como ejercicio demostrar que si $x_1, x_2, x_3, \ldots$ e $y_1, y_2, y_3, \ldots$ son las sucesiones del Ejemplo 1, que aproximan a $\sqrt{2}$ y π, respectivamente, entonces, como $\lim_{k \to \infty} x_k = \sqrt{2}$ y $\lim_{k \to \infty} y_k = \pi$, se tiene que $\lim_{k \to \infty} (x_k, y_k) = (\sqrt{2}, \pi)$.

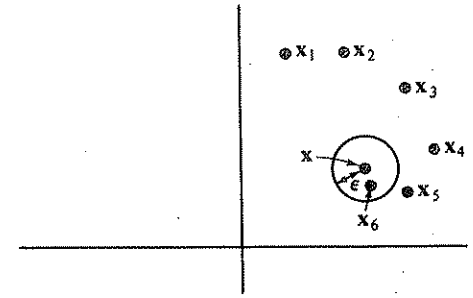


Figura 40

Consideremos ahora el método de Newton para aproximar una solución de la ecuación $f(x) = 0$ en el caso en que f es real y x es una variable real. Si sucede que el gráfico de f es convexo, como en el caso de la Fig. 41, entonces es geométricamente claro que la recta tangente al gráfico $(x_0, f(x_0))$ corta el eje x en un punto x_1, que es una aproximación mejor que x_0 de la solución $\bar{x}$. Una vez elegido x_0 en forma más o menos arbitraria y encontrado x_1, el proceso

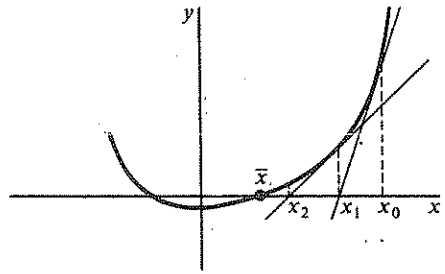


Figura 41

puede ser repetido. Esta vez usamos la recta tangente en $(x_1, f(x_1))$ y denotamos por x_2 la intersección de esta recta con el eje x. Obtenemos así una sucesión $x_0, x_1, x_2, \ldots$ de números que aproximan a $\bar{x}$.

En la práctica, necesitamos una fórmula para calcular la sucesión $x_1, x_2, \ldots$ Notemos primero que la recta tangente en $(x_0, f(x_0))$ tiene ecuación

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Como la aproximación x_1 se obtiene intersectando la tangente con el eje x, hacemos $y = 0$ en la ecuación y despejamos x_1. El resultado es

$$0 = f'(x_0)(x_1 - x_0) + f(x_0),$$

de donde

$$f'(x_0)(x_1 - x_0) = -f(x_0).$$

Si $f'(x_0) \neq 0$,

$$x_1 - x_0 = -\frac{f(x_0)}{f'(x_0)}$$

o bien

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Para encontrar x_2 una vez que se tiene x_1, basta remplazar x_0 por x_1 en esta última ecuación, obteniéndose

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

En general, tenemos la siguiente fórmula para x_{k+1},

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Ejemplo 2. La ecuación $x^2 - 3 = 0$ tiene dos soluciones, $\sqrt{3}$ y $-\sqrt{3}$. Para aproximar $\sqrt{3}$, elegimos $x_0 = 2$ y calculamos x_{k+1} a partir de x_k usando la Ec. (1) que, en este caso, es

$$x_{k+1} = x_k - \frac{(x_k^2 - 3)}{2x_k}$$

$$= \frac{(x_k^2 + 3)}{2x_k}.$$

Por tanto, obtenemos $x_1 = \frac{7}{4} = 1{,}75$. Sustituyendo este valor en la fórmula con $k = 1$, nos da $x_2 = \frac{97}{56} \approx 1{,}7321$. Esta aproximación de $\sqrt{3}$ es válida hasta tres decimales, hecho que no ha sido demostrado.

Podemos usar un método análogo al recién descrito en el caso en que f es una función de $\mathcal{R}^n$ en $\mathcal{R}^n$. La diferencia está en que en este caso $\mathbf{x}_1$, $\mathbf{x}_0$ y $f(\mathbf{x}_0)$ son vectores en $\mathcal{R}^n$ y $f'(\mathbf{x}_0)$ es una matriz jacobiana n-por-n. Para aproximar una solución de la ecuación vectorial $f(\mathbf{x}_0)$, consideramos la ecuación que define el valor de la mejor aproximación afín de f cerca de $\mathbf{x}_0$; es decir,

$$\mathbf{y} = f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x}_0), \tag{2}$$

donde $\mathbf{x}_0$ es una aproximación inicial de la solución deseada $\bar{\mathbf{x}}$. En la Fig. 41, la Ec. (2) puede ser interpretada como la ecuación de la tangente al gráfico de f en $(\mathbf{x}_0, f(\mathbf{x}_0))$. Como antes, hacemos $\mathbf{y} = 0$ en la Ec. (2), obteniendo

$$0 = f'(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0) + f(\mathbf{x}_0),$$

de donde

$$f'(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0) = -f(\mathbf{x}_0).$$

Ahora bien, si $f'(\mathbf{x}_0)$ tiene una matriz inversa $[f'(\mathbf{x}_0)]^{-1}$, podemos tomar los inversos de ambos miembros, deduciendo que

$$\mathbf{x}_1 - \mathbf{x}_0 = -[f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}_0).$$

Finalmente,

$$\mathbf{x}_1 = \mathbf{x}_0 - [f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}_0).$$

En esta ecuación $[f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}_0)$ es el vector obtenido aplicando la inversa de la matriz $f'(\mathbf{x}_0)$ al vector $f(\mathbf{x}_0)$. El vector $\mathbf{x}_1$ es una primera mejora de la primera aproximación $\mathbf{x}_0$ de la solución $\bar{\mathbf{x}}$.

Lo que hemos hecho puede ser repetido remplazando $\mathbf{x}_0$ por $\mathbf{x}_1$, con lo cual se tiene

$$\mathbf{x}_2 = \mathbf{x}_1 - [f'(\mathbf{x}_1)]^{-1} f(\mathbf{x}_0).$$

Luego de $k + 1$ pasos, tenemos

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [f'(\mathbf{x}_k)]^{-1} f(\mathbf{x}_k). \tag{9.1}$$

Ejemplo 3. Consideremos el par de ecuaciones

$$x^2 + y^2 = 2$$
$$x^2 - y^2 = 1.$$

Para encontrar una solución aproximada por el método de Newton, definimos

$$f(x, y) = \begin{pmatrix} x^2 + y^2 - 2 \\ x^2 - y^2 - 1 \end{pmatrix}$$

y resolvemos la ecuación $f(x, y) = 0$. Es claro que f es una función de $\mathcal{R}^2$ en $\mathcal{R}^2$. Como exigimos que $x^2 + y^2 - 2 = 0$ y $x^2 - y^2 - 1 = 0$, es útil dibujar las curvas definidas por estas ecuaciones. En la Fig. 42, las soluciones exactas

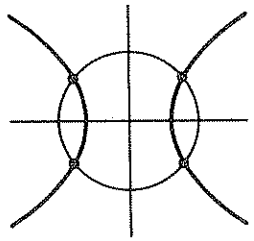


Figura 42

están representadas por los cuatro puntos en que se intersectan el círculo $x^2 + y^2 - 2 = 0$ y la hipérbola $x^2 - y^2 - 1 = 0$. La elección de una aproximación inicial depende de la solución que queremos aproximar. Para encontrar la solución del primer cuadrante, probamos con $\mathbf{x}_0 = (1, 1)$. Como

$$f(x, y) = \begin{pmatrix} x^2 + y^2 - 2 \\ x^2 - y^2 - 1 \end{pmatrix},$$

tenemos que

$$f'(x, y) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$$

y

$$[f'(x, y)]^{-1} = \begin{pmatrix} \frac{1}{4}x^{-1} & \frac{1}{4}x^{-1} \\ \frac{1}{4}y^{-1} & -\frac{1}{4}y^{-1} \end{pmatrix}.$$

Por tanto,

$$\begin{pmatrix} x \\ y \end{pmatrix} - [f'(x, y)]^{-1} f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \frac{1}{4}x^{-1} & \frac{1}{4}x^{-1} \\ \frac{1}{4}y^{-1} & -\frac{1}{4}y^{-1} \end{pmatrix} \begin{pmatrix} x^2 + y^2 - 2 \\ x^2 - y^2 - 1 \end{pmatrix}$$

$$= \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \dfrac{2x^2 - 3}{4x} \\ \dfrac{2y^2 - 1}{4y} \end{pmatrix} = \begin{pmatrix} \dfrac{2x^2 + 3}{4x} \\ \dfrac{2y^2 + 1}{4y} \end{pmatrix}. \tag{3}$$

Este vector corresponde a la expresión $(x^2 + 2)/2x$ del ejemplo anterior y es la fórmula para encontrar la sucesión de aproximaciones. Tomando $\mathbf{x}_0 = (x_0, y_0) = (1, 1)$, obtenemos

$$\mathbf{x}_1 = \begin{pmatrix} \dfrac{2x_0^2 \quad 3}{4x} \\ \dfrac{2y_0^2 + 1}{4y_0} \end{pmatrix} = \begin{pmatrix} \frac{5}{4} \\ \frac{3}{4} \end{pmatrix} = \begin{pmatrix} 1{,}25 \\ 0{,}75 \end{pmatrix}.$$

Sustituyendo $\mathbf{x}_1$ en (1), tenemos

$$\mathbf{x}_2 = \begin{pmatrix} \dfrac{2(1{,}25)^2 + 3}{4 \quad 25)} \\ \dfrac{2(0{,}75) + 3}{4(0{,}75)} \end{pmatrix} \approx \begin{pmatrix} 1{,}225 \\ 1{,}70833 \end{pmatrix}.$$

Sustituyendo ahora $\mathbf{x}_2$, tenemos

$$\mathbf{x}_3 = \begin{pmatrix} \dfrac{2(1{,}225)^2 + 3}{4(1{,}225)} \\ \dfrac{2(1{,}70833)^2 + 3}{4(1{,}70833)} \end{pmatrix} \approx \begin{pmatrix} 1{,}22574 \\ 0{,}707108 \end{pmatrix}.$$

Análogamente,

$$\mathbf{x}_4 \approx \begin{pmatrix} 1{,}22474 \\ 0{,}707107 \end{pmatrix}$$

y

$$\mathbf{x}_5 \approx \begin{pmatrix} 1{,}22474 \\ 0{,}707107 \end{pmatrix}.$$

Como en el caso anterior, si continuamos con solo cinco decimales, no hay cambios.

En este ejemplo es posible resolver el sistema por eliminación, obteniéndose

$\mathbf{x} = (\sqrt{1{,}5}, \sqrt{0{,}5})$. La aproximación $\mathbf{x}_4 = (1{,}22474, 0{,}707107)$ es correcta hasta el quinto decimal. Los otros tres vectores solución pueden ser encontrados por simetría. Considerando la Fig. 39, vemos que los podemos obtener cambiando el signo de una o ambas coordenadas. El método numérico también podría haber sido aplicado a partir de $(-1, -1)$, $(-1, 1)$ o $(1, -1)$.

Es necesario tener cuidado al elegir la aproximación inicial $\mathbf{x}_0$, de modo que ella sea suficientemente buena. Por ejemplo, si en el Ejemplo 3 queremos la solución en el primer cuadrante, es claro que un error demasiado grosero en la elección de $\mathbf{x}_0$ nos llevaría a la solución errada. En muchos ejemplos, un dibujo o un análisis geométrico de la función f indica cómo elegir $\mathbf{x}_0$.

Al usar el método de Newton en dimensiones elevadas, puede ser muy tedioso invertir la matriz $f'(\mathbf{x}_n)$ en cada paso de la iteración. Entonces se puede usar la fórmula de Newton modificada:

9.2
$$\mathbf{x}_{k+1} = \mathbf{x}_k - [f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}_k)$$

para obtener una sucesión de aproximaciones de la solución de $f(x) = 0$. Para $k = 0$ esta fórmula coincide con la fórmula de Newton. Para $k \geq 1$, $\mathbf{x}_{k+1}$, definido por la Ec. 8.2, será en general distinto del valor determinado por la fórmula de Newton, ya que la matriz $[f'(\mathbf{x}_0)]^{-1}$ no varía en cada paso de la Ec. 8.2.

Ejemplo 4. Volviendo a la ecuación del Ejemplo 3, es decir,

$$\begin{pmatrix} x^2 + y^2 - 2 \\ x^2 - y^2 - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

aplicamos la Ec. 9.2 con $\mathbf{x}_0 = (1, 1)$. Entonces

$$[f'(\mathbf{x}_0)]^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} \end{pmatrix},$$

de modo que

$$\mathbf{x} - [f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}) = \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} x^2 + y^2 - 2 \\ x^2 - y^2 - 1 \end{pmatrix}$$

$$= \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \dfrac{2x^2 - 3}{4} \\ \dfrac{2y^2 - 1}{4} \end{pmatrix} = \begin{pmatrix} \dfrac{-2x^2 + 4x + 3}{4} \\ \dfrac{-2y^2 + 4y + 1}{4} \end{pmatrix}.$$

Entonces $\mathbf{x}_1$, definido como $\mathbf{x}_1 = \mathbf{x}_0 - [f'(\mathbf{x}_0)]^{-1} f(\mathbf{x}_0)$ es

$$\mathbf{x}_1 = \begin{pmatrix} \dfrac{-2x_0^2 + 4x_0 + 3}{4} \\ \dfrac{2y_0^2 + 4y_0 + 1}{4} \end{pmatrix} = \begin{pmatrix} 1{,}25 \\ 0{,}75 \end{pmatrix}.$$

En el paso siguiente,

$$\mathbf{x}_2 = \begin{pmatrix} \dfrac{-2(1\ 25)^2 + 4(1{,}25) + 3}{4} \\ \dfrac{-2(0{,}75)^2 + 4(0{,}75) + 1}{4} \end{pmatrix} = \begin{pmatrix} 1{,}21875 \\ 0{,}71875 \end{pmatrix}$$

Continuando así, llegamos a

$$\mathbf{x}_{10} = \begin{pmatrix} 1{,}22474 \\ 0{,}707107 \end{pmatrix},$$

que es igual al resultado del Ejemplo 3, pero que ha sido obtenido en un número mayor de etapas.

Al decidir si se usa la fórmula de Newton 9.1 o su modificación 9.2, dos hechos deben ser tenidos en cuenta. En general, la Ec. 9.1 converge más rápidamente que 9.2; es decir, se tiene un error más pequeño en un determinado número de pasos. Por otra parte, la Ec. 9.2 tiene la ventaja de que solo requiere el cálculo de la inversa de una matriz. Por tanto, si el cálculo de las matrices inversas $[f'(\mathbf{x}_k)]^{-1}$ es demasiado largo, podrá valer la pena efectuar los pasos extra que requiere la aplicación de 9.2. En la Sec. 5 del Apéndice demostraremos el Teorema 5.1, que garantiza la convergencia de la sucesión en 9.2 bajo ciertas condiciones. El Teorema 5.1 se usa también para demostrar el teorema de la función inversa, considerado en la Sec. 6 del Cap. 4.

EJERCICIOS

1. (a) Dibuje el gráfico de $f(x) = \sqrt[3]{x} - x$ para $-2 < x < 2$.
   (b) Dibuje las rectas tangentes al gráfico de f en $x_0 = \frac{3}{4}$ y $x_0 = -\frac{1}{4}$.
   (c) Para cada uno de los x_0 de la parte (b), ¿a qué solución de $f(x) = 0$ se puede esperar que converja la iteración de Newton?
   (d) Discuta la elección de $x_0 = \sqrt{3}/9$ como primera aproximación para la solución de $f(x) = 0$.
2. Demuestre que el método de Newton no produce convergencia a la solución única de $\sqrt[3]{x} = 0$ a no ser que se tome el valor inicial $x_0 = 0$. [*Indicación*: Demuestre que $x_k = (-2)^k x_0$.]
3. (a) Demuestre que si se elige x_0 entonces para aproximar la solución de $\cos x - x = 0$, para todo $k \geq 0$, se tiene

$$x_{k+1} = \frac{x_k \operatorname{sen} x_k + \cos x_k}{1 + \operatorname{sen} x_k}.$$

(b) Use un computador para calcular x_{10}.

4. Encuentre soluciones aproximadas del par de ecuaciones

$$x^2 + y - 1 = 0$$
$$x + y^2 - 2 = 0$$

en los siguientes pasos:

(a) Dibuje las curvas representadas por cada ecuación.
(b) Defina f por

$$f(x, y) = \begin{pmatrix} x^2 + y - 1 \\ x + y^2 - 2 \end{pmatrix},$$

encuentre $f'(x, y)$, $[f'(x, y)]^{-1}$ y $(x, y) - [f'(x, y)]^{-1} f(x, y)$.
(c) Usando el dibujo de la parte (a) elija una solución aproximada $\mathbf{x}_0 = (x_0, y_0)$ de la solución de $f(x, y) = 0$ en el cuarto cuadrante del plano xy.
(d) Calcule $\mathbf{x}_1 = (x_1, y_1)$ usando la Ec. 9.1.
(e) Use un computador para encontrar $\mathbf{x}_5 = (x_5, y_5)$.

5. Sea

$$f(x, y, z) = \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 \\ x^3 + y^3 + z^3 \end{pmatrix}$$

Encuentre $f'(1, 1, -1)$. Tomando $\mathbf{x}_0 = (1, 2, -1)$ aplique la fórmula de Newton modificada 9.2 para aproximar la solución de $f(x, y, z) = (2, 6, 8)$.

6. Sea

$$g(u, v) = \begin{pmatrix} u^2 + uv^2 \\ u + v^3 \end{pmatrix}.$$

Note que $g(1, 1) = (2, 2)$ y use el método de Newton de la Ec. 9.1 o su modificación 9.2, para aproximar la solución de $g(u, v) = (1, 9, 2, 1)$.

4

Cálculo vectorial

1. Derivadas direccionales

Una derivada parcial de una función real mide la velocidad de cambio de la función en la dirección de una coordenada. Para medir la velocidad de cambio en una dirección arbitraria, usamos la derivada direccional. Recordemos que un **vector unitario u** es un vector de largo 1, es decir, tal que $|\mathbf{u}| = 1$. Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ una función real y sea $\mathbf{u}$ un vector unitario en el espacio dominio $\mathcal{R}^n$. La **derivada de f con respecto a u**, denotada por $\partial f/\partial \mathbf{u}$, es la función real definida por

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x})}{t}.$$

El dominio $\partial f/\partial \mathbf{u}$ es el subconjunto del dominio de f formado por los puntos $\mathbf{x}$ tales que el límite existe.

El teorema que sigue expresa la relación que existe entre la derivada con respecto a un vector unitario y la diferencial.

1.1 Teorema

Si f es diferenciable en $\mathbf{x}$, entonces

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{u}$$

para todo vector unitario $\mathbf{u}$ en $\mathcal{R}^n$.

Demostración. La existencia de la derivada $f'(\mathbf{x})$ implica que

$$\lim_{t\to 0}\frac{f(\mathbf{x}+t\mathbf{u})-f(\mathbf{x})-f'(\mathbf{x})(t\mathbf{u})}{|t\mathbf{u}|}=0,$$

lo cual es equivalente a

$$\lim_{t\to 0}\frac{1}{|\mathbf{u}|}\left|\frac{f(\mathbf{x}+t\mathbf{u})-f(\mathbf{x})}{t}-f'(\mathbf{x})\mathbf{u}\right|=0.$$

Lo cual, a su vez, es equivalente a

$$\lim_{t\to 0}\frac{f(\mathbf{x}+\mathbf{u})-f(\mathbf{x})}{t}=f'(\mathbf{x})\mathbf{u},$$

y la demostración queda terminada.

La ecuación

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x})=f'(\mathbf{x})\mathbf{u}$$

muestra que para una función diferenciable, la derivada direccional no incluye nada realmente nuevo. Sin embargo, ella proporciona una importante interpretación de la derivada $f'(\mathbf{x})$ aplicada al vector unitario $\mathbf{u}$.

Para todo vector $\mathbf{u}$ en $\mathcal{R}^n$, de largo $|\mathbf{u}|=1$ hemos definido la **derivada direccional de f en la dirección de $\mathbf{u}$** como la función $\partial f/\partial \mathbf{u}$. Este nombre se justifica, ya que, en el espacio euclidiano existe una manera natural de asociar un vector a cada dirección que es la de tomar el vector unitario en esa dirección. El número $(\partial f/\partial \mathbf{u})(\mathbf{x})$ es, entonces la manera estándar de medir el cambio de los valores de f en la dirección de $\mathbf{u}$.

Ejemplo 1. El espacio dominio de la función

$$f(x,y,z)=xyz+e^{2x+y}$$

es el espacio tridimensional. Calculemos la derivada direccional de f en la dirección de $\mathbf{u}=(\frac{1}{2},\frac{1}{2},1/\sqrt{2})$. Tomando $\mathbf{x}=(x,y,z)$ y usando el Teorema 1.1, obtenemos

$$\frac{df}{\partial \mathbf{u}}(\mathbf{x})=f'(\mathbf{x})\mathbf{u}=\begin{pmatrix}yz+2e^{2x+y} & xz+e^{2x+y} & xy\end{pmatrix}\begin{pmatrix}1/2\\1/2\\1/\sqrt{2}\end{pmatrix}$$

$$=\frac{yz+xz+\sqrt{2}xy}{2}+\frac{3e^{2x+y}}{2}.$$

Por tanto, la derivada direccional de f en la dirección de $\mathbf{u}$ tiene, en el origen, el valor $\partial f/\partial \mathbf{u}\ (0,0,0)=\frac{3}{2}$.

Sea $\mathcal{R}^2\xrightarrow{f}\mathcal{R}$ una función cuyo gráfico es una superficie en el espacio euclidiano tridimensional y sea $\mathbf{u}$ un vector unitario en $\mathcal{R}^2$, es decir, tal que $|\mathbf{u}|=1$.

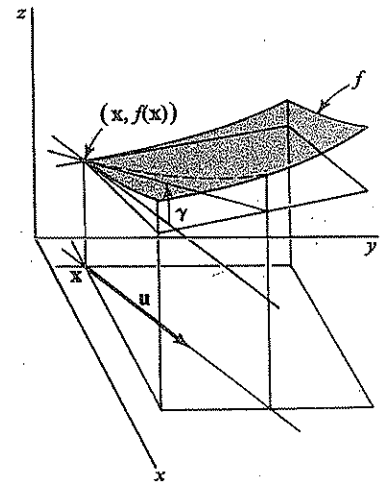


Figura 1

La Fig. 1 muestra un ejemplo. El valor de la derivada direccional $\partial f/\partial \mathbf{u}$ en $\mathbf{x}=(x,y)$ es, por definición,

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x})=\lim_{t\to 0}\frac{f(\mathbf{x}+t\mathbf{u})-f(\mathbf{x})}{t}.$$

La distancia entre los puntos $\mathbf{x}+t\mathbf{u}$ y $\mathbf{x}$ está dada por

$$|(\mathbf{x}+t\mathbf{u})-\mathbf{x}|=|t\mathbf{u}|=|t|.$$

Por tanto, el cociente

$$\frac{f(\mathbf{x}+t\mathbf{u})-f(\mathbf{x})}{t}$$

es la pendiente de la recta que pasa por los puntos $f(\mathbf{x}+t\mathbf{u})$ y $f(\mathbf{x})$. De esto sigue que, el límite $\partial f/\partial \mathbf{u}$ de este cociente es la pendiente de la recta tangente, en $(\mathbf{x},f(\mathbf{x}))$, a la curva formada por la intersección del gráfico de f con el plano paralelo al eje z que contiene a $\mathbf{x}$ y $\mathbf{x}+\mathbf{u}$. En la Fig. 1, esta curva ha sido representada por una línea punteada. El ángulo γ de la Fig. 1 satisface la ecuación

$$\tan\gamma=\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}).$$

Esta situación es una generalización de la situación ilustrada en la Fig. 24 de

la Sec. 5 del Cap. 5. Si tomamos $\mathbf{u} = (1\ 0)$, el ángulo γ es el ángulo α de esa figura y

$$\frac{\partial f}{\partial \mathbf{u}} = \frac{\partial f}{\partial x}.$$

Por otra parte, si tomamos $\mathbf{v} = (0, 1)$, entonces γ es el ángulo β de la Fig. 24, y

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{\partial f}{\partial y}.$$

El teorema del valor medio para funciones reales de una variable real puede ser extendido para funciones *reales* de un vector variable como sigue:

1.2 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ una función diferenciable en un conjunto abierto que contiene al segmento de recta S que une los vectores $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$. Entonces, existe un punto $\mathbf{x}_0$ en S tal que

$$f(\mathbf{y}) - f(\mathbf{x}) = f'(\mathbf{x}_0)(\mathbf{y} - \mathbf{x}).$$

Demostración. Consideremos la función $g(t) = f(t(\mathbf{y} - \mathbf{x}) + \mathbf{x})$, definida para $0 \leq t \leq 1$. Hagamos $m(t) = t(\mathbf{y} - \mathbf{x}) + \mathbf{x}$, entonces, si h es un número real,

$$\begin{aligned} g(t + h) - g(t) &= f\big((t + h)(\mathbf{y} - \mathbf{x}) + \mathbf{x}\big) - f\big(t(\mathbf{y} - \mathbf{x}) + \mathbf{x}\big) \\ &= f'\big(m(t)\big)h(\mathbf{y} - \mathbf{x}) + |h(\mathbf{y} - \mathbf{x})|\, Z\big(h(\mathbf{y} - \mathbf{x})\big), \end{aligned}$$

donde $\lim_{h \to 0} Z(h(\mathbf{y} - \mathbf{x})) = 0$. Esta relación dice simplemente que f es diferenciable en $m(t)$. Dividiendo por h, tenemos

$$\frac{g(t + h) - g(t)}{h} = f\big(m(t)\big)(\mathbf{y} - \mathbf{x}) \pm |\mathbf{y} - \mathbf{x}|\, Z\big(h(\mathbf{y} - \mathbf{x})\big),$$

de donde

$$g'(t) = f'\big(m(t)\big)(\mathbf{y} - \mathbf{x}). \tag{1}$$

Pero, aplicando el teorema del valor medio para funciones de una variable,

$$\begin{aligned} \frac{g(1) - g(0)}{1 - 0} &= f(\mathbf{y}) - f(\mathbf{x}) \\ &= g'(t_0), \end{aligned}$$

para algún t_0 que satisface $0 < t_0 < t_1$. Haciendo $t = t_0$ y $m(t_0) = \mathbf{x}_0$ en la Ec. (1) obtenemos el resultado buscado.

Una de las conclusiones más importantes que pueden ser deducidas del teorema del valor medio para funciones de una variable es que una función

cuya derivada es cero en un intervalo, es constante en ese intervalo. Para una función f de una variable vectorial, remplazamos el intervalo dominio por un conjunto abierto D en $\mathcal{R}^n$ que supondremos es **poligonalmente conexo**. Un conjunto S es poligonalmente conexo si todo par de puntos en él pueden ser unidos por una poligonal en S, es decir, por un número finito de segmentos de recta en S.

1.3 Teorema

Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es diferenciable en un conjunto abierto y poligonalmente conexo D satisface $f'(\mathbf{x}) = 0$ para todo $\mathbf{x}$ en D, entonces f es constante.

Demostración. Basta demostrar que cada una de las funciones coordenadas de f es constante y, por tanto, podemos suponer que f es real. Si $\mathbf{x}_1$ y $\mathbf{x}_2$ son puntos de D que pueden ser unidos por un segmento de recta en D, entonces el Teorema 1.2 y la hipótesis que dice que $f'(\mathbf{x}) = 0$ en D, implican que $f(\mathbf{x}_1) = f(\mathbf{x}_2)$. Es evidente que esta conclusión también es válida para dos puntos $\mathbf{x}_1$ y $\mathbf{x}_2$ unidos por una sucesión finita de segmentos de recta. Por tanto, f es constante.

EJERCICIOS

1. Para cada una de las siguientes funciones definidas en el espacio euclidiano de tres dimensiones, encuentre la derivada direccional en la dirección del vector unitario $\mathbf{u}$ en el punto $\mathbf{x}$.

   (a) $f(x, y, z) = x^2 + y^2 + z^2$, $\mathbf{u} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, $\mathbf{x} = (1, 0, 1)$.
   (b) $h(x, y, z) = xyz$, $\mathbf{u} = (\cos\alpha \operatorname{sen}\beta, \operatorname{sen}\alpha \operatorname{sen}\beta, \cos\beta)$, $\mathbf{x} = (1, 0, 0)$.
   (c) $f(x, y)$, $\mathbf{x} = (x, y)$, y $\mathbf{y} = (\cos\alpha, \operatorname{sen}\alpha)$. (Suponer que f es real y diferenciable.)

2. Para cada una de las siguientes funciones reales definidas en el plano euclidiano, encuentre la derivada direccional en $\mathbf{x}$ y en la dirección indicada.

   (a) $f(x, y) = x^2 - y^2$ en $\mathbf{x} = (1, 1)$ y en la dirección

   $$\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right). \qquad \left[\textit{Resp}: \frac{-2}{\sqrt{5}}.\right]$$

   (b) $f(x, y) = e^x \operatorname{sen} y$ en $\mathbf{x} = (1, 0)$ y en la dirección de $(\cos\alpha, \operatorname{sen}\alpha)$.
   (c) $f(x, y) = e^{x+y}$ en $\mathbf{x} = (1, 1)$ y en la dirección de la curva definida por $g(t) = (t^2, t^3)$ en $g(2)$, para t aumentando.

3. Encuentre el valor absoluto de la derivada direccional en $(1, 1, 0)$ de la función $f(x, y, z) = x^2 + ye^z$ tomada en la dirección de la recta tangente en $g(0)$ a la curva en $\mathcal{R}^3$ definida paramétricamente por $g(t) = (3t^2 + t + 1, 2t, t^2)$.

4. Encuentre la derivada direccional en $(1, 0, 0)$ de la función $f(x, y, z) = x^2 + ye^z$ en la dirección en que aumenta t a lo largo de la curva $g(t) = (t^2 - t + 2, t, t + 2)$ en el punto $g(0)$. $[\textit{Resp}.\ -1/\sqrt{3}.]$

5. Encuentre el valor absoluto de la derivada direccional en $(1, 0, 1)$ de la función $f(x, y, z) = 4x^2 + y^2z$ en la dirección de la perpendicular en $(1, 1, 1)$ a la superficie en $\mathcal{R}^3$ definida implícitamente por $x^2 + 2y^2 + z^2 = 4$. [*Resp.* $8/\sqrt{6}$.]

6. (a) Demuestre que el vector $(y_1z_2 - y_2z_1, z_1x_2 - z_2x_1, x_1y_2 - x_2y_1)$ es perpendicular a (x_1, y_1, z_1) y a (x_2, y_2, z_2).
   (b) Encuentre el valor absoluto de la derivada direccional en $(1, 2, 1)$ de la función $f(x, y, z) = x^3 + y^2 + z$ en la dirección de la perpendicular en $(1, 2, 1)$ a la superficie definida paramétricamente por

$$(x, y, z) = (u^2v, u + v, u). \qquad [\textit{Resp}: 2/\sqrt{3}.]$$

7. Si la temperatura de una esfera sólida de radio 3 con entero en $(0, 0, 0)$ está dada por $T(x, y, z) = yz + zx + xy$ en el punto (x, y, z), encuentre la dirección en la cual la temperatura aumenta más rápidamente en el punto $(1, 1, 2)$.

8. Demuestre que la fórmula del valor medio dada en el Teorema 1.2 puede ser escrita en la forma

$$\frac{f(\mathbf{y}) - f(\mathbf{x})}{|\mathbf{y} - \mathbf{x}|} = \frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0),$$

donde $\mathbf{u} = (\mathbf{y} - \mathbf{x})/|\mathbf{y} - \mathbf{x}|$.

9. Demuestre que la función f definida por

$$f(x, y) = \begin{cases} \dfrac{x\,|y|}{\sqrt{x^2 + y^2}}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

tiene derivada direccional en toda dirección en el punto $(0, 0)$, pero que f no es diferenciable en $(0, 0)$.

2. El gradiente

En la primera sección del Cap. 1, consideramos algunos ejemplos de campos vectoriales. Muchos de los más importantes aparecen como campos gradientes que serán descritos más adelante.

Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ es una función real diferenciable, entonces la función ∇f, definida por

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right)$$

se llama el **gradiente** de f. Evidentemente, el gradiente es una función de $\mathcal{R}^n$ en $\mathcal{R}^n$ y, generalmente, se considera como un campo vectorial.

Ejemplo 1. La función $f(x, y) = \frac{1}{6}(x^2 + y^3)$ es diferenciable en todo $\mathcal{R}^2$, de modo que $\nabla f(x, y) = (\frac{1}{3}x, \frac{1}{2}y^2)$ está también definida en todo $\mathcal{R}^2$. La Fig. 2 muestra a este campo en varios puntos.

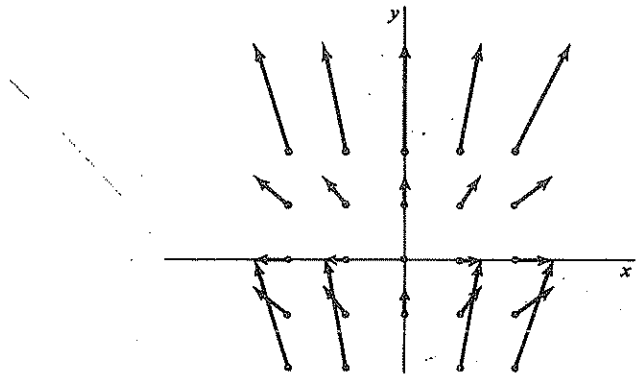


Figura 2

La función $g(x, y, z) = x^2 + y^2 + z^2$ tiene gradiente $\nabla g(x, y, z) = (2x, 2y, 2z)$ y la dirección del campo está en la dirección desde el origen al punto.

El gradiente de una función es importante por varias razones. En primer lugar, notemos que la derivada direccional de una función real f con respecto a un vector $\mathbf{u}$ unitario puede ser expresada en términos del gradiente de f en la forma

$$\textbf{2.1} \qquad \frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u}.$$

Esto se debe a que $(\partial f/\partial \mathbf{u})(\mathbf{x}) = f'(\mathbf{x})\mathbf{u}$ y el resultado de la aplicación de la matriz $f'(\mathbf{x})$ a $\mathbf{u}$ es el producto punto $\nabla f(\mathbf{x}) \cdot \mathbf{u}$. Usando la Ec. 2.1, podemos demostrar fácilmente el teorema que sigue. (Este es el origen del uso del término gradiente que generalmente significa inclinación o pendiente.)

2.2 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ diferenciable en un conjunto abierto D en $\mathcal{R}^n$. Entonces, en todo punto $\mathbf{x}$ en D tal que $\nabla f(\mathbf{x}) \neq 0$, el vector $\nabla f(\mathbf{x})$ tiene la dirección del aumento máximo de f. El número $|\nabla f(\mathbf{x})|$ es la velocidad máxima de aumento de f.

Demostración. Dado un vector unitario $\mathbf{u}$, tenemos que, por la Ec. 2.1 y la desigualdad de Cauchy-Schwarz,

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} \leq |\nabla f(\mathbf{x})|\,|\mathbf{u}| = |\nabla f(\mathbf{x})|$$

Pero cuando $\mathbf{u} = \nabla f(\mathbf{x})/|\nabla f(\mathbf{x})|$, y solo entonces, tenemos que

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}) = |\nabla f(\mathbf{x})|.$$

Por tanto, la velocidad de aumento $\partial f/\partial \mathbf{u}(\mathbf{x})$ no es jamás mayor que $|\nabla f(\mathbf{x})|$ y es igual en la dirección del gradiente.

Ejemplo 2. Sea $f(x, y) = e^{xy}$. Entonces $\nabla f(x, y) = (ye^{xy}, xe^{yx})$, de modo que en $(1, 2)$ la función f aumenta con rapidez máxima en la dirección $\nabla f(1, 2) = (2e^2, e^2)$ que tiene la misma dirección que el vector unitario $(2/\sqrt{5}, 1/\sqrt{5})$. La velocidad de aumento en esa dirección es $|\nabla f(1, 2)| = \sqrt{5}\, e^2$. En forma análoga, $\nabla f(-1, 2) = (2e^{-2}, -e^{-2})$ tiene la dirección de $(2/\sqrt{5}, -1/\sqrt{5})$, con velocidad máxima de aumento en $(-1, 2)$ igual a $\sqrt{5}\, e^{-2}$. La velocidad máxima de disminución está en la dirección opuesta, a saber, $(-2/\sqrt{5}, 1/\sqrt{5})$.

A continuación demostraremos una regla de cadena para la derivación de la composición $g(f(t))$ de una función $\mathfrak{R} \xrightarrow{f} \mathfrak{R}^n$ y una función $\mathfrak{R}^n \xrightarrow{g} \mathfrak{R}$. Así, por ejemplo, si $\mathfrak{R} \xrightarrow{f} \mathfrak{R}^3$ está definida por $f(t) = (t, t^2, t)$ y $\mathfrak{R}^3 \xrightarrow{g} \mathfrak{R}$ por $g(x, y, z) = x \cos(y + z)$, entonces $g(f(t)) = t \cos(t^2 + t)$. Esto define una función de $\mathfrak{R}$ en $\mathfrak{R}$. Por ejemplo, si g está definida en una región D en $\mathfrak{R}^3$ y f describe el movimiento de un punto a lo largo de un camino contenido en D, podemos estar interesados en encontrar la velocidad de cambio de la función compuesta, con respecto de t. El teorema da una fórmula para hacer esto en términos del gradiente de g.

2.3 Teorema

Sea g una función real continuamente diferenciable en un conjunto abierto D en $\mathfrak{R}^n$ y $f(t)$ definida y diferenciable para $a < t < b$ y con valores en D. Entonces la función compuesta $F(t) = g(f(t))$ es diferenciable para $a < t < b$ y

$$F'(t) = \nabla g(f(t)) \cdot f'(t).$$

Demostración. Por definición,

$$F'(t) = \lim_{h \to 0} \frac{F(t + h) - F(t)}{h}$$

$$= \lim_{h \to 0} \frac{g(f(t + h)) - g(f(t))}{h},$$

si este límite existe. Como f es diferenciable, ella es continua. Entonces podemos encontrar $\delta > 0$ tal que cuando $|h| < \delta$, $f(t + h)$ está siempre dentro de una bola abierta de centro en $f(t)$ y contenida en D. Ahora, aplicamos a la función g el teorema del valor medio de la sección anterior, obtenemos

$$g(\mathbf{y}) - g(\mathbf{x}) = g'(\mathbf{x}_0)(\mathbf{y} - \mathbf{x})$$

$$= \nabla g(\mathbf{x}_0) \cdot (\mathbf{y} - \mathbf{x}),$$

donde $\mathbf{x}_0$ es algún punto en el segmento de recta que une a y y x. Tomando $\mathbf{x} = f(t)$ y $\mathbf{y} = f(t + h)$, donde $|h| < \delta$, tenemos

$$\frac{F(t + h) - F(t)}{h} = \nabla g(\mathbf{x}_0) \cdot \frac{f(t + h) - f(t)}{h}.$$

Ahora el vector $\mathbf{x}_0$ es algún punto en el segmento de recta que une a $f(t)$ con $f(t + h)$. (¿Por qué está $\mathbf{x}_0$ en el dominio de f?) Como supusimos que g es continuamente diferenciable, $\nabla g(\mathbf{x})$ es continua, de modo que $\nabla g(\mathbf{x}_0)$ tiende a $\nabla g(f(t))$ cuando h tiende a cero. El producto punto es continuo, de modo que $F'(t)$ existe y

$$F'(t) = \lim_{h \to 0} \nabla g(\mathbf{x}_0) \cdot \frac{f(t + h) - f(t)}{h}$$

$$= \nabla g(f(t)) \cdot f'(t).$$

Ejemplo 3. Sea $g(x, y) = x^2y + xy^3$ para todo (x, y) en $\mathfrak{R}^2$. Sea $f(t)$ diferenciable en una vecindad de $t = t_0$ y con valores en $\mathfrak{R}^2$. Solo se sabe que $f(t_0) = (-1, 1)$ y $f'(t_0) = (2, 3)$, de modo que solo se conoce la función compuesta $F(t) = g(f(t))$ en $t = t_0$ y $F'(t_0)$ no puede ser calculada por derivación directa. Sin embargo, por el teorema anterior, tenemos

$$F'(t_0) = \nabla g(f(t_0)) \cdot f'(t_0).$$

Tenemos que $\nabla g(x, y) = (2xy + y^3, x^2 + 3xy^2)$, de modo que $\nabla g(F(t_0)) = (-1, -2)$. Entonces $F(t_0) = (-1, -2) \cdot (2, 3) = -8$.

En la próxima sección demostraremos una extensión del Teorema 2.3 al caso en que f toma valores vectoriales.

El gradiente es especialmente útil para analizar los conjuntos de nivel de una función real. Recordemos que un conjunto de nivel S de una función f es el conjunto de los puntos $\mathbf{x}$ que satisfacen $f(\mathbf{x}) = k$ para alguna constante k. Para $\mathfrak{R}^2 \xrightarrow{f} \mathfrak{R}$, generalmente estamos interesados en S cuando él es una curva y, para $\mathfrak{R}^3 \xrightarrow{f} \mathfrak{R}$, los conjuntos S que se consideran generalmente son las superficies. La Fig. 3 muestra algunos ejemplos.

Decir que un punto $\mathbf{x}_0$ está en el conjunto de nivel S que corresponde a k es lo mismo que decir que $f(\mathbf{x}_0) = k$. Ahora bien, supongamos que existe una curva lisa γ en S que pasa por $\mathbf{x}_0$ y con vector tangente $\mathbf{v}$ en $\mathbf{x}_0$. Parametrizamos γ mediante $\mathfrak{R} \xrightarrow{g} \mathfrak{R}^n$, con $g(t_0) = \mathbf{x}_0$ y $g'(t_0) = \mathbf{v} \neq 0$. La Fig. 3 muestra esta situación. Aplicando la regla de la cadena a la función $h(t) = f(g(t))$ en t_0, tenemos que

$$h'(t_0) = \nabla f(g(t_0)) \cdot g'(t_0)$$

$$= \nabla f(\mathbf{x}_0) \cdot \mathbf{v}.$$

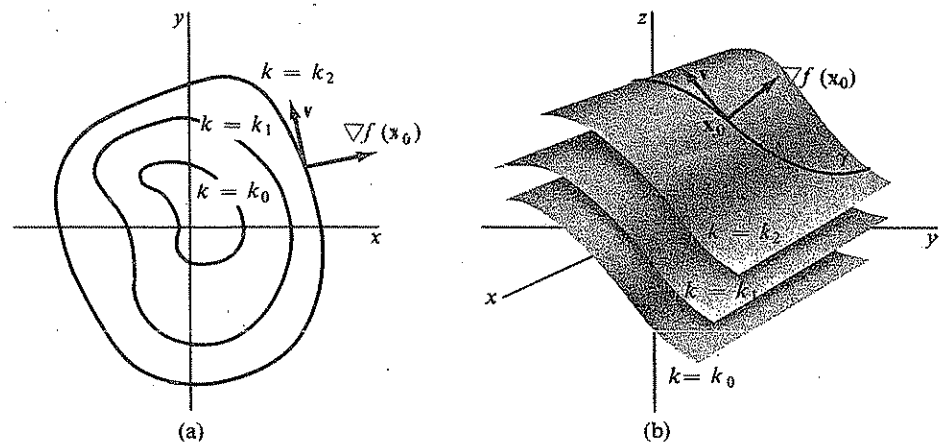


Figura 3

Pero, como γ está en S, tenemos que $h(t) = f(g(t)) = k$, es decir, h es constante. Por tanto, $h'(t_0) = 0$ y

$$\nabla f(\mathbf{x}_0) \cdot \mathbf{v} = 0.$$

Esto indica que $\nabla f(\mathbf{x}_0)$, si no es cero, es perpendicular a todo vector tangente a una curva lisa arbitraria en S y que pasa por $\mathbf{x}_0$. Por esta razón es natural decir que $\nabla f(\mathbf{x}_0)$ es perpendicular o **normal** a S en $\mathbf{x}_0$ y tomar como **plano** (o recta) **tangente** a S en $\mathbf{x}_0$ al conjunto de todos los puntos $\mathbf{x}$ que satisfacen

$$\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0, \qquad \text{si} \quad \nabla f(\mathbf{x}_0) \neq 0.$$

Ejemplo 4. Una de las superficies de nivel de la función $f(x, y, z) = x^2 + y^2 - z^2$ es el cono C formado por todos los puntos que satisfacen $x^2 + y^2 - z^2 = 0$. El punto $\mathbf{x}_0 = (1, 1, \sqrt{2})$ está en C y, para encontrar el plano tangente a C en $\mathbf{x}_0$, calculamos que $\nabla f(\mathbf{x}_0) = (2, 2, -2\sqrt{2})$. Entonces

$$\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = (2, 2, -2\sqrt{2}) \cdot (x - 1, y - 1, z - \sqrt{2}),$$

y el plano tangente está dado por $(x - 1) + (y - 1) - \sqrt{2}(z - \sqrt{2}) = 0$ o $x + y - \sqrt{2}\, z = 0$. La Fig. 4 muestra este plano. Nótese que C y su tangente contienen una recta común con dirección $(1, 1, \sqrt{2})$ y que el vector normal a la tangente es perpendicular a esa recta.

En vista del Teorema 2.2 y de 2.4, vemos que la dirección de aumento máximo en un punto de una función real diferenciable, es perpendicular al conjunto de nivel de la función que pasa por ese punto.

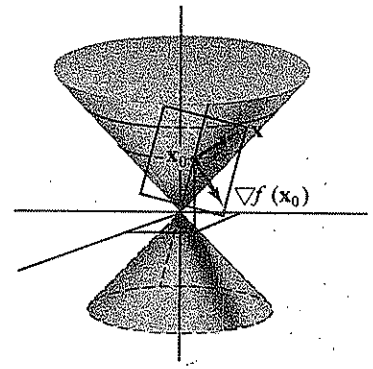


Figura 4

Ejemplo 5. La función $f(x, y) = xy$ tiene curvas de nivel $xy = k$. Si $k \neq 0$, estas curvas son hipérbolas y cada una de ellas es intersectada perpendicularmente por cada uno de los miembros de la familia de hipérbolas $g(x, y) = x^2 - y^2 = k$. (La Fig. 5 muestra estas curvas.) Para comprobar esto, notemos que $\nabla f(x, y) = (y, x)$ y $\nabla g(x, y) = (2x, -2y)$. Por tanto, para cada (x, y), tenemos que $\nabla f(x, y) \cdot \nabla g(x, y) = 0$. Luego los vectores normales y, por tanto, también los tangentes, son perpendiculares. También se deduce que toda tangente a una curva de una familia tiene la dirección del aumento máximo de la función que define a la otra familia. El argumento falla en (0, 0). Véase Ejercicio 6.

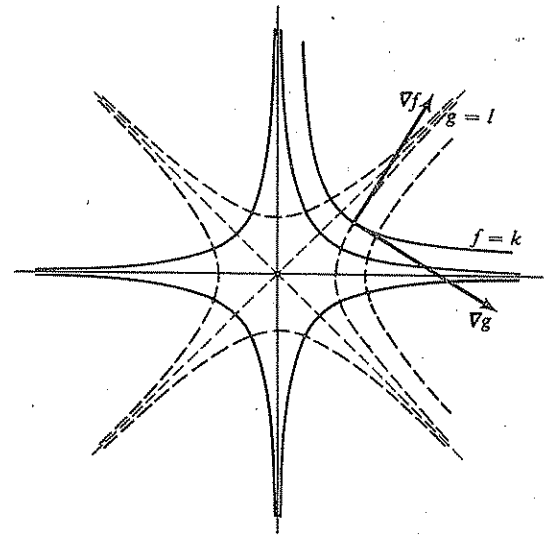


Figura 5

Como un nuevo ejemplo del uso del gradiente, demostraremos el teorema siguiente, que es una generalización de la fórmula bien conocida

$$\int_a^b f'(t)\,dt = f(b) - f(a), \tag{1}$$

llamada el teorema fundamental del cálculo.

2.5 Teorema

Sea f una función real continuamente diferenciable definida en un conjunto abierto D en $\mathcal{R}^n$. (Por tanto, ∇f es un campo vectorial continuo en D.) Si γ es una curva lisa por pedazos en D con puntos inicial y final $\mathbf{a}$ y $\mathbf{b}$, respectivamente, entonces

$$\int_\gamma \nabla f \cdot d\mathbf{x} = f(\mathbf{b}) - f(\mathbf{a}).$$

En particular, el valor de la integral de línea de un campo *gradiente* a lo largo de una curva depende solo de los puntos extremos de la curva y, por tanto, en este caso, se justifica la notación

$$\int_a^b \nabla f \cdot d\mathbf{x} = f(\mathbf{b}) - f(\mathbf{a})$$

Demostración. Supongamos que γ está parametrizada por $G(t)$ con $a \le t \le b$ y que $G(a) = \mathbf{a}$, $G(b) = \mathbf{b}$. Usando primero la definición de la integral de línea y luego el Teorema 2.3, tenemos que

$$\int_\gamma \nabla f \cdot d\mathbf{x} = \int_a^b \nabla f\big(G(t)\big) \cdot G'(t)\,dt$$

$$= \int_a^b \frac{d}{dt} f\big(G(t)\big)\,dt.$$

Pero por la Ec. (1) (el teorema fundamental del cálculo), esta última integral es igual a $f(G(b)) - f(G(a)) = f(\mathbf{b}) - f(\mathbf{a})$.

Ejemplo 6. Consideremos el campo $\nabla f(x, y)$ en $\mathcal{R}^2$, donde $f(x, y) = \frac{1}{2}(x^2 + y^2)$. Entonces $\nabla f(x, y) = (x, y)$. Si γ es una curva continuamente diferenciable con puntos inicial y final (x_1, y_1) y (x_2, y_2), entonces

$$\int_\gamma x\,dx + y\,dy = f(x_2, y_2) - f(x_1, y_1)$$

$$= \tfrac{1}{2}(x_2^2 - x_1^2) + \tfrac{1}{2}(y_2^2 - y_1^2).$$

Esto es lo que podemos esperar formalmente del teorema fundamental. Por otra parte, si tomamos $g(x, y) = xy$, entonces tenemos que $\nabla g(x, y) = (y, x)$ y, para cualquier curva como la que recientemente consideramos, tenemos que

$$\int_\gamma y\,dx + x\,dy = x_2y_2 - x_1y_1.$$

EJERCICIOS

1. Encuentre el gradiente ∇f para cada una de las funciones en los puntos indicados.
   (a) $f(x, y, z) = (x - y)e^{xz}$; $(1, 2, -1)$.
   (b) $f = x^2 - y^2 = \operatorname{sen} y$, para (x, y) arbitrario en $\mathcal{R}^2$.
   (c) $f(x, y) = x + y$; $(2, 3)$.
   (d) $f(\mathbf{x}) = |\mathbf{x}|^2$ para $\mathbf{x}$ arbitrario en $\mathcal{R}^n$.
   (e) $f(\mathbf{x}) = |\mathbf{x}|^\alpha$ para $\mathbf{x} \neq 0$ arbitrario en $\mathcal{R}^n$.
   (f) $f(x, y) = 0$ idénticamente en $\mathcal{R}^2$, $(1, 1)$.
2. Calcule la dirección y velocidad de aumento máximo para cada una de las funciones del problema anterior, en el punto indicado.
3. (a) La notación grad f se usa a menudo para el gradiente ∇f. Demuestre que
   $$\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}) = \operatorname{grad} f(\mathbf{x}) \cdot \mathbf{y}.$$
   (b) A menudo se usa la notación $\nabla_{\mathbf{y}} f$ para la derivada $\partial f/\partial \mathbf{y}$. Demuestre que
   $$\nabla_{\mathbf{y}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{y}.$$
4. Si es posible, encuentre el vector normal y el plano tangente a cada una de las curvas o superficies de nivel en los puntos indicados.
   (a) $x^2 + y^2 - z^2 = 2$ en $(x, y, z) = (1, 1, 0)$ y en $(x, y, z) = (0, 0, 0)$.
   (b) $x \operatorname{sen} y = 0$ en $(x, y) = (0, \pi/2)$ y en $(x, y) = (0, 0)$.
   (c) $|\mathbf{x}| = 2$ en $\mathbf{x} = \mathbf{e}_1$, el primer vector de la base natural de $\mathcal{R}^n$.
   (d) $x^2y + yz + w = 3$ en $(x, y, z, w) = (1, 1, 1, 1)$.
   (e) $xyz = 1$ en $(x, y, z) = (1, 1, 1)$.
   (f) $xyz = 0$ en $(x, y, z) = (1, 2, 0)$.
5. Si $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$ es continuamente diferenciable, su gráfico puede ser definido implícitamente en $\mathcal{R}^3$ como la superficie de nivel S de la función $F(x, y, z) = z - f(x, y)$ dada por $F(x, y, z) = 0$.
   (a) Demuestre que $\nabla F = (-\partial f/\partial x, -\partial f/\partial y, 1)$, vector que nunca es cero.
   (b) Encuentre el vector normal y el plano tangente al gráfico de $f(x, y) = xy + ye^x$ en $(x, y) = (1, 1)$.
6. (a) El ejemplo $f(x, y) = x^2 + y^2$ tiene $\nabla f(0, 0) = 0$, lo cual no indica si existe una dirección de aumento máximo para f en $(x, y) = (0, 0)$. ¿Es esto razonable? ¿Qué sucede en $(0, 0)$?
   (b) En el Ejemplo 5 de la Sec. 2, el punto $(x, y) = (0, 0)$ no ha sido considerado.

¿Cuáles son las direcciones de aumento máximo para $f(x, y) = xy$ y $g(x, y) = x^2 - y^2$ en $(x, y) = (0, 0)$?

7. Si $f(x, y) = e^{x+y}$ y $F(0) = (1, 2)$, encuentre $g'(0)$, donde $g(t) = f(F(t))$ y $F(0) = (1, 1)$.

8. Sea γ una curva en $\mathbb{R}^3$ recorrida con rapidez 2 y dirección $(1, -1, 2)$ en el instante $t = 1$. Si a $t = 1$ corresponde el punto $(1, 1, 1)$ en γ, encuentre la velocidad de cambio de la función $x + y + xy$ a lo largo de γ en $t = 1$.

9. Si $f(x, y, z) = \operatorname{sen} x$ y $F(t) = (\cos t, \operatorname{sen} t, t)$, encuentre $g'(\pi)$, donde $g(t) = f(F(t))$.

10. Sea $\mathbb{R} \xrightarrow{F} \mathbb{R}^n$ diferenciable. Sea $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ continuamente diferenciable tal que la composición $g(t) = f(F(t))$ existe. Si $F'(t_0)$ es tangente a la superficie de nivel de f en $F(t_0)$, demuestre que $g'(t_0) = 0$.

11. Dado un campo vectorial $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^n$, el problema de encontrar una función $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ tal que $\nabla f = F$ es equivalente a la solución del siguiente sistema de ecuaciones en f:

$$\frac{\partial f}{\partial x_1} = F_1, \ldots, \frac{\partial f}{\partial x_n} = F_n,$$

donde $F_1, \ldots, F_n$ son las funciones coordenadas de F.

(a) en el caso $n = 2$, demuestre que si el sistema

$$\frac{\partial f}{\partial x}(x, y) = F_1(x, y), \qquad \frac{\partial f}{\partial y}(x, y) = F_2(x, y)$$

tiene una solución f, entonces f debe tener las dos formas

$$f(x, y) = \int F_1(x, y)\, dx + C_1(y)$$

y

$$f(x, y) = \int F_2(x, y)\, dy + C_2(x),$$

donde cada una de las integrales indefinidas se hace manteniendo la otra variable fija.

(b) Usando la parte (a), encuentre f si $\nabla f(x, y) = (y^2 + 2xy, 2xy + x^2)$.

(c) Usando una extensión apropiada de la parte (a), encuentre f tal que $f(x, y, z) = (y + z, z + x, x + y)$.

(d) Encuentre f tal que $\nabla f(x, y, z) = (yz + z, xz, xy + x)$.

12. (a) Encuentre la función f del Prob. 11(b) calculando una integral de línea de ∇f de $(0, 0)$ a (x, y).

(b) Encuentre la función f del Prob. 11(c) calculando una integral de línea de $(0, 0, 0)$ a (x, y, z).

13. (a) Demuestre que si $F_1, \ldots, F_n$ son continuamente diferenciables, entonces una condición necesaria para que el sistema

$$\frac{\partial f}{\partial x_1} = F_1, \ldots, \frac{\partial f}{\partial x_n} = F_n$$

tenga una solución f es que

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}, \qquad i, j = 1, \ldots, n.$$

(b) Demuestre que las funciones $\mathbb{R}^2 \xrightarrow{F} \mathbb{R}^2$ y $\mathbb{R}^3 \xrightarrow{G} \mathbb{R}^3$ definidas por $F(x, y) = (xy, x^2)$ y $G(x, y, z) = (y, x, -zx)$ no son gradientes de funciones reales.

14. (a) Calcule la integral de línea $\int_\gamma y\, dx + x\, dy$ a lo largo de una curva continuamente diferenciable arbitraria de $(0, 0)$ a $(2, 3)$. [*Indicación*: Adivine una función $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}$ tal que $\nabla f(x, y) = (y, x)$.]

(b) Calcule la integral de línea $\int_\gamma x\, dx + y\, dy + z\, dz$ a lo largo de una curva continuamente diferenciable arbitraria entre $(0, 0, 0)$ y $(1, 2, 3)$ y de $(0, 0, 0)$ a (x, y, z).

15. Sea $F(x, y) = \left(\dfrac{-y}{x^2 + y^2}, \dfrac{x}{x^2 + y^2} \right)$ para $(x, y) \neq (0, 0)$.

(a) Si F_1 y F_2 son las funciones coordenadas de F, demuestre que

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}.$$

(b) Demuestre que, en cualquier región que no interseque al eje y, F es el gradiente de $f(x, y) = \operatorname{arctg}(y/x)$, en su rama principal.

(c) Demuestre que F no es el gradiente de una función diferenciable f en una región que engloba al origen —por ejemplo, la región definida por $x^2 + y^2 > 0$.

16. Demuestre que si un campo de fuerzas en una región D de $\mathbb{R}^3$ tiene la forma ∇f para alguna función continuamente diferenciable $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}$ entonces el trabajo realizado al mover una partícula en el campo solo depende de las superficies de nivel de f en las cuales la partícula está al principio y al final del movimiento.

17. Las superficies de nivel de una función $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}$ se llaman las **superficies equipotenciales** del campo vectorial ∇f y f se llama la **función potencial** del campo.

(a) Demuestre que las superficies equipotenciales son ortogonales al campo.

(b) Encuentre las superficies equipotenciales del campo $\nabla f(x, y, z) = (x, y, z)$.

(c) Encuentre el campo para el cual $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$ (el **potencial newtoniano**) es la función potencial.

(d) Encuentre el campo del cual $f(x, y) = -\frac{1}{2} \log (x^2 + y^2)$ (el **potencial logarítmico**) es la función potencial.

(e) Demuestre que si $f(\mathbf{x}) = |\mathbf{x}|^{2-n}$ (el **potencial newtoniano generalizado** en $\mathbb{R}^n$, $n \geq 3$), entonces $\nabla f(\mathbf{x}) = (2 - n)\, |\mathbf{x}|^{-n} \mathbf{x}$.

18. Sea γ una curva lisa definida por $\mathbb{R} \xrightarrow{F} \mathbb{R}^3$ para $a \leq t \leq b$ tal que está en una región D de $\mathbb{R}^3$ en la cual $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}$ es continuamente diferenciable. Demuestre que

$$\frac{d}{dt} \int_{F(a)}^{F(t)} \nabla f \cdot d\mathbf{x} = \nabla f(F(t)) \cdot F'(t).$$

19. Suponga que $\int_{x_0}^{x} F(\mathbf{x}) \cdot d\mathbf{x}$ es independiente de la curva lisa por pedazos γ que une a $\mathbf{x}_0$ para todo $\mathbf{x}$ en algún conjunto abierto D en $\mathbb{R}^n$. Demuestre que $\nabla f = F$, donde $f(\mathbf{x}) = \int_{x_0}^{x} F \cdot d\mathbf{x}$. [*Indicación*: Considere a la integral a lo largo de una curva γ que

se aproxima a $\mathbf{x}$ en una dirección arbitraria y aplique el teorema fundamental del cálculo.]

20. Si $T(x, y, z)$ representa la temperatura en el punto (x, y, z) de una región R en $\mathcal{R}^3$, el campo vectorial ∇T se llama el **gradiente de temperatura.** Bajo ciertas suposiciones físicas, $\nabla T(x, y, z)$ es proporcional al vector que representa la dirección y razón por unidad de área del flujo calórico en (x, y, z). Los conjuntos en los cuales T es constante se llaman isotermas. Si las isotermas de una función de temperatura son esferas concéntricas, demuestre que el gradiente de temperatura está dirigido ya sea hacia el centro de las esferas ya sea en el sentido opuesto.

3. La regla de la cadena

Una de las fórmulas más útiles del cálculo de una variable es la regla de la cadena que se usa para calcular la derivada de la composición de una función con otra:

$$[g(f(x))]' = g'(f(x))f'(x).$$

La generalización al caso de varias variables es igualmente útil y, si se la formula en forma adecuada, es igualmente fácil de expresar.

Si dos funciones f y g están relacionadas de modo que el espacio imagen de f es igual al espacio dominio de g, entonces podemos formar la **función compuesta** $g \circ f$ aplicando primero f y después g. Es decir, para todo vector $\mathbf{x}$ en el dominio de f y tal que $f(\mathbf{x})$ está en el dominio de g, tenemos que

$$g \circ f(\mathbf{x}) = g(f(\mathbf{x}))$$

El dominio de $g \circ f$ está formado por los vectores $\mathbf{x}$ que son llevados por f al dominio g. La Fig. 6 da una idea de la composición de funciones.

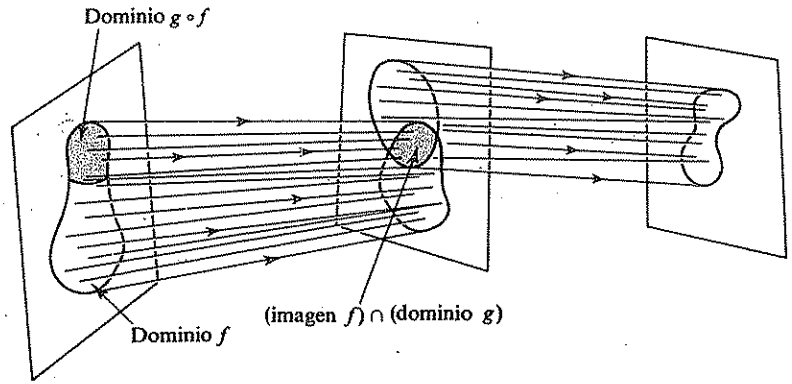


Figura 6

Ejemplo 1. Supongamos dada una región de dimensión 2 en que puede moverse un punto de acuerdo con una ley dada. Se sabe que si un punto parte de la posición inicial (u, v), entonces, después de un cierto tiempo especificado, el punto se encuentra en la posición (x, y). Estas posiciones están dadas por ecuaciones de la forma

$$x = g_1(u, v)$$

$$y = g_2(u, v).$$

En notación vectorial, estas ecuaciones tienen la forma

$$\mathbf{x} = g(\mathbf{u}),$$

donde $\mathbf{x} = (x, y)$, $\mathbf{u} = (u, v)$ y g tiene funciones coordenadas g_1, g_2. Ahora supongamos que la posición inicial $\mathbf{u} = (u, v)$ está determinada, en función de otras variables (s, t), por las ecuaciones

$$u = f_1(s, t)$$

$$v = f_2(s, t).$$

Estas ecuaciones pueden ser escritas en forma vectorial como

$$\mathbf{u} = f(\mathbf{s}),$$

donde $\mathbf{s} = (s, t)$ y f tiene funciones coordenadas f_1, f_2. Entonces (x, y) y (s, t) están relacionadas por

$$x = g_1(f_1(s, t), f_2(s, t))$$

$$y = g_2(f_1(s, t), f_2(s, t)),$$

es decir,

$$\mathbf{x} = g(f(\mathbf{s})).$$

Si usamos la notación $g \circ f$ para la compuesta de g y f, también podemos escribir

$$\mathbf{x} = g \circ f(\mathbf{s}).$$

Para determinar la derivada de $g \circ f$ en términos de las derivadas de g y f, suponemos que $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es diferenciable en $\mathbf{x}_0$ y que $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^p$ es diferenciable en $\mathbf{y}_0 = f(\mathbf{x}_0)$. Entonces $g'(\mathbf{y}_0)$ es una matriz p-por-m y $f'(\mathbf{x}_0)$ una matriz m-por-n. Por tanto, el producto $g'(\mathbf{y}_0)f'(\mathbf{x}_0)$ está definido y es una matriz p-por-n. La regla de la cadena dice que esta matriz producto es la derivada de $g \circ f$ en $\mathbf{x}_0$. Como la matriz producto corresponde a la composición de funciones lineales, este resultado puede ser expresado en términos de diferenciales como sigue: la diferencial de una composición es la composición de las diferenciales.

Ejemplo 2. Consideramos el caso especial en que f es una función de una variable real y g es real. Entonces $g \circ f$ es una función real de una variable real. El Teorema 2.3 muestra que si f y g son continuamente diferenciables, entonces

$$(g \circ f)'(t) = \nabla g(f(t)) \cdot f'(t).$$

Es decir, en términos de las funciones coordenadas,

$$(g \circ f)'(t) = \left(\frac{\partial g}{\partial y_1}(f(t)), \ldots, \frac{\partial g}{\partial y_m}(f(t))\right) \cdot (f_1'(t), \ldots, f_n'(t)).$$

El término de la derecha de esta última ecuación puede ser expresado como producto matricial en términos de las matrices derivadas.

$$g'(f(t)) = \left(\frac{\partial g}{\partial y_1}(f(t)), \ldots, \frac{\partial g}{\partial y_m}(f(t))\right),$$

y

$$f'(t) = \begin{pmatrix} f_1'(t) \\ \cdot \\ \cdot \\ \cdot \\ f_m'(t) \end{pmatrix}$$

ya que $(g \circ f)'(t) = g'(f(t))f'(t)$. El producto de $g'(f(t))$ y $f'(t)$ es la multiplicación matricial de una matriz 1-por-m y una matriz m-por-1 y es equivalente al producto punto de las dos matrices consideradas como vectores en $\mathcal{R}^m$. Por tanto, en el caso en que tanto el dominio de f como la imagen de g son de dimensión 1, las fórmulas

$$\nabla g(f(t)) \cdot f'(t) \quad \text{y} \quad g'(f(t))f'(t)$$

son prácticamente lo mismo.

El teorema que daremos es una extensión al caso de dimensiones arbitrarias para el dominio e imagen de f y g.

3.1 La regla de la cadena

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ continuamente diferenciable en $\mathbf{x}$ y $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^p$ continuamente diferenciable en $f(\mathbf{x})$. Si $g \circ f$ está definida en un conjunto abierto que contiene a x, entonces $g \circ f$ es continuamente diferenciable en $\mathbf{x}$ y

$$(g \circ f)'(\mathbf{x}) = g'(f(\mathbf{x}))f'(\mathbf{x}).$$

Demostración. Solo necesitamos demostrar que la matriz derivada de $g \circ f$ en $\mathbf{x}$ tiene elementos que dependen continuamente de los elementos del producto de $g'(f(\mathbf{x}))$ y $f'(\mathbf{x})$. Estas matrices tienen la forma

$$\begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\mathbf{x})) & \cdots & \frac{\partial g_1}{\partial y_m}(f(\mathbf{x})) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \frac{\partial g_p}{\partial y_1}(f(\mathbf{x})) & \cdots & \frac{\partial g_p}{\partial y_m}(f(\mathbf{x})) \end{pmatrix} \text{ y } \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_m}(\mathbf{x}) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_m}{\partial x_m}(\mathbf{x}) \end{pmatrix}.$$

El producto de las matrices tienen como elemento ij la suma de productos

$$\sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(f(\mathbf{x}))\frac{\partial f_k}{\partial x_j}(\mathbf{x}). \tag{1}$$

Pero esta expresión es simplemente el producto punto de los vectores $\Delta g_i(f(\mathbf{x}))$ y $\partial f/\partial x_j(\mathbf{x})$. Por tanto, por el Teorema 2.3, tenemos que

$$\nabla g_i(f(\mathbf{x})) \cdot \frac{\partial f}{\partial x_j}(\mathbf{x}) = \frac{\partial (g_i \circ f)}{\partial x_j}(\mathbf{x}), \tag{2}$$

ya que estamos derivando con respecto de la variable x_j. Pero esto establece la relación matricial, ya que, por definición, los elementos en $(g \circ f)'(\mathbf{x})$ están dados por el lado derecho de la Ec. (2). Como g y f son continuamente diferenciables, la Ec. (1) representa una función continua en $\mathbf{x}$ para cada i y j. Por tanto, $g \circ f$ es continuamente diferenciable.

En la Sec. 2 del Apéndice daremos una versión de la regla de la cadena en que solo suponemos que f y g son diferenciables y concluimos que $g \circ f$ es diferenciable (no necesariamente continuamente diferenciable).

Ejemplo 3. Sea $f(x, y) = (x^2 + y^2, x^2 - y^2)$ y sea $g(u, v) = (uv, u + v)$. Calculamos

$$g'(u, v) = \begin{pmatrix} v & u \\ 1 & 1 \end{pmatrix} \quad \text{y} \quad f'(x, y) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}.$$

Para calcular $(g \circ f)'(2, 1)$, notamos que $f(2, 1) = (5, 3)$ y obtenemos

$$g'(5, 3) = \begin{pmatrix} 3 & 5 \\ 1 & 1 \end{pmatrix} \quad \text{y} \quad f'(2, 1) = \begin{pmatrix} 4 & 2 \\ 4 & -2 \end{pmatrix}.$$

Entonces el producto de estas matrices da

$$(g \circ f)'(2, 1) = \begin{pmatrix} 32 & -4 \\ 8 & 0 \end{pmatrix}.$$

En cálculo, es muy práctico denotar una función por el mismo símbolo que un elemento típico de su imagen. De esta manera, la derivada de una función

$\mathcal{R} \xrightarrow{f} \mathcal{R}$ dada por la ecuación $y = f(x)$, se denota por dy/dx. Del mismo modo, las derivadas parciales de una función $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}$ se denotan por

$$\frac{\partial w}{\partial x}, \qquad \frac{\partial w}{\partial y}, \qquad \text{y} \qquad \frac{\partial w}{\partial z}$$

junto con la ecuación aclaratoria $w = f(x, y, z)$. Por ejemplo, si $w = xy^2e^{x+3z}$, entonces

$$\frac{\partial w}{\partial x} = y^2e^{x+3z} + xy^2e^{x+3z},$$

$$\frac{\partial w}{\partial y} = 2xye^{x+3z},$$

$$\frac{\partial w}{\partial z} = 3xy^2e^{x+3z}.$$

Esta notación tiene la desventaja de que no contiene referencias específicas a la función que se está derivando. Por otra parte, ésta es una notación muy conveniente y es el lenguaje tradicional del cálculo. Para ilustrar esta conveniencia, supongamos que f y g están dadas por

$$w = g(x, y, z), \qquad x = f_1(s, t), \qquad y = f_2(s, t), \qquad z = f_3(s, t).$$

Entonces, por la regla de la cadena,

$$\begin{pmatrix} \frac{\partial w}{\partial s} & \frac{\partial w}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \\ \frac{\partial z}{\partial s} & \frac{\partial z}{\partial t} \end{pmatrix}.$$

Efectuando la multiplicación matricial, se tiene

$$\left.\begin{aligned} \frac{\partial w}{\partial s} &= \frac{\partial g}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial s} \\ \frac{\partial w}{\partial t} &= \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial \partial}\frac{\partial z}{\partial t} \end{aligned}\right\}. \tag{3}$$

Si el dominio de f es 1-dimensional, es decir, si f es función de una variable, la aplicación de la regla de la cadena toma una forma aparentemente distinta. Por ejemplo, consideremos

$$w = g(u, v), \qquad \begin{pmatrix} u \\ v \end{pmatrix} = f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}.$$

La compuesta $g \circ f$ es, en este caso, una función real de una variable. Su diferencial está definida por una matriz 1-por-1 cuyo único elemento es la derivada

$$\frac{d(g \circ f)}{dt} = \frac{dw}{dt}.$$

Las derivadas de g y f están definidas por matrices jacobianas

$$\begin{pmatrix} \frac{\partial w}{\partial u} & \frac{\partial w}{\partial v} \end{pmatrix} \qquad \text{y} \qquad \begin{pmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{pmatrix},$$

respectivamente. Por tanto, la regla de la cadena implica que

$$\frac{dw}{dt} = \begin{pmatrix} \frac{\partial w}{\partial u} & \frac{\partial w}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{pmatrix} = \frac{\partial w}{\partial u}\frac{du}{dt} + \frac{\partial w}{\partial v}\frac{dv}{dt}. \tag{4}$$

Finalmente, supongamos que f y g son funciones reales de una variable. Esta es la situación del cálculo en una variable. La derivada de f en t, la de g en $s = f(t)$ y la de $g \circ f$ en t están representadas por tres matrices 1-por-1, $(f'(t))$, $(g'(s))$ y $(g \circ f)'(t)$, respectivamente. La regla de la cadena implica que

3.2 $$(g \circ f)'(t) = g'(s)f'(t).$$

Si las funciones están representadas en la forma

$$x = g(s), \qquad s = f(t),$$

la famosa ecuación

$$\frac{dx}{dt} = \frac{dx}{ds}\frac{ds}{dt}. \tag{5}$$

es una forma más explícita de la Ec. 3.2.

Ejemplo 4. Si

$$\begin{cases} x = u^2 + v^3 \\ y = e^{uv}, \end{cases} \qquad \text{y} \qquad \begin{cases} u = t + 1 \\ v = e^t, \end{cases}$$

encontraremos dx/dt y dy/dt en $t = 0$. Sea $\mathbb{R} \xrightarrow{f} \mathbb{R}^2$ y $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$ las funciones definidas por

$$f(t) = \begin{pmatrix} t+1 \\ e^t \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad -\infty < t < \infty,$$

$$g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 + v^3 \\ e^{uv} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{matrix} -\infty < u < \infty. \\ -\infty < v < \infty. \end{matrix}$$

La diferencial de f en t está dada por la matriz jacobiana 2-por-1

$$\begin{pmatrix} \dfrac{du}{dt} \\ \dfrac{dv}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ e^t \end{pmatrix}.$$

La matriz de la diferencial de g en $\begin{pmatrix} u \\ v \end{pmatrix}$ es

$$\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} 2u & 3v^2 \\ ve^{uv} & ue^{uv} \end{pmatrix}.$$

La dependencia de x e y de t está dada por

$$\begin{pmatrix} x \\ y \end{pmatrix} = (g \circ f)(t), \qquad -\infty < t < \infty.$$

Por tanto, las dos derivadas dx/dt y dy/dt son los elementos de la matriz jacobiana que define la diferencial de la función compuesta $g \circ f$. Por tanto, la regla de la cadena implica que

$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix} \begin{pmatrix} \dfrac{du}{dt} \\ \dfrac{dv}{dt} \end{pmatrix}.$$

Es decir,

$$\left.\begin{aligned} \frac{dx}{dt} &= \frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt} = 2u + 3v^2e^t \\ \frac{dy}{dt} &= \frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt} = ve^{uv} + ue^{uv+t} \end{aligned}\right\}. \tag{6}$$

Si $t = 0$, entonces $\begin{pmatrix} u \\ v \end{pmatrix} = f(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, y obtenemos que $u = v = 1$. Por tanto,

$$\frac{dx}{dt}(0) = 2 + 3 = 5,$$

$$\frac{dy}{dt}(0) = e + e + 2e.$$

La definición de la multiplicación matricial hace fácil la memorización de las fórmulas que resultan de la aplicación de la regla de la cadena. La pauta general de estas fórmulas queda clara cuando se denotan las funciones coordenadas como variables reales, como lo hicimos en las Ecs. (3), (4), (5) y (6). Todas las fórmulas de la forma general

$$\ldots + \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} + \ldots$$

tienen la desventaja de no contener referencias específicas a los puntos en los cuales se deben evaluar las derivadas. Por supuesto, es esencial tener esta información que se puede obtener de la fórmula

$$(g \circ f)'(\mathbf{x}) = g'(f(\mathbf{x}))f'(\mathbf{x})$$

Por tanto, las derivadas que aparecen en la matriz $f'(\mathbf{x})$ se evalúan en $\mathbf{x}$ y las de la matriz $g'(f(\mathbf{x}))$, en $f(\mathbf{x})$. Esta es la razón por la cual hicimos $t = 0$ y $u = v = 1$ en la Ec. (6), para obtener el resultado final del Ejemplo 4.

Ejemplo 5. Sea

$$z = xy \quad \text{y} \quad \begin{matrix} x = f(u, v). \\ y = g(u, v). \end{matrix}$$

Supongamos que, para $u = 1$ y $v = 0$, tenemos que

$$\frac{\partial x}{\partial u} = -1, \qquad \frac{\partial x}{\partial v} = 3, \qquad \frac{\partial y}{\partial u} = 5, \qquad \frac{\partial y}{\partial v} = 0.$$

Supongamos también que $f(1, 2) = 2$ y $g(1, 2) = -2$. Calculamos $(\partial z/\partial u)(1, 2)$. La regla de la cadena implica que

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}. \tag{7}$$

Cuando $u = 1$ y $v = 2$, tenemos que $x = f(1, 2) = 2$ e $y = g(1, 2) = -2$. Por tanto,

$$\frac{\partial z}{\partial x}(2, -2) = y|_{x=2, y=-2} = -2$$

$$\frac{\partial z}{\partial y}(2, -2) = x|_{x=2, y=-2} = 2.$$

Para calcular $\partial z/\partial u$ en $(u, v) = (1, 2)$, es necesario saber en qué puntos evaluar las derivadas parciales que aparecen en la Ec. (7). Más detalladamente, la regla de la cadena implica que

$$\frac{\partial z}{\partial u}(1, 2) = \frac{\partial z}{\partial x}(2, -2)\frac{\partial x}{\partial u}(1, 2) + \frac{\partial z}{\partial y}(2, -2)\frac{\partial y}{\partial u}(1, 2).$$

Por tanto,

$$\frac{\partial z}{\partial u}(1, 2) = (-2)(-1) + (2)(5) = 12.$$

Ejemplo 6. Sea $w = f(ax^2 + bxy + cy^2)$ e $y = x^2 + x + 1$. Calcularemos $dw/dx(-1)$. Para esto, usamos las fórmulas que se deducen de la regla de la cadena [tales como (3), (4), (5) y (7)]. Hagamos

$$z = ax^2 + bxy + cy^2.$$

Entonces, $w = f(z)$ y

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}.$$

Por tanto,

$$\begin{aligned}\frac{dw}{dx} &= \frac{df}{dz}\frac{dz}{dx} = \frac{df}{dz}\left(\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}\right)\\ &= f'(z)\big(2ax + by + (bx + 2cy)(2x + 1)\big).\end{aligned}$$

Si $x = -1$, entonces $y = 1$, de modo que $z = a - b + c$. Por tanto,

$$\frac{dw}{dx}(-1) = f'(a - b + c)(-2a + 2b - 2c).$$

La matriz jacobiana o derivada de una función f de $\mathcal{R}^n$ en $\mathcal{R}^n$ es cuadrada y, por tanto, tiene un determinante. Este determinante, $\det f'(x)$, es una función real de x, llamada el determinante jacobiano de f. Este determinante juega un papel muy importante en el teorema del cambio de variables para la integral que veremos en el Cap. 6. En este punto notamos un corolario sencillo de la regla de la cadena y de la fórmula producto para determinantes:

3.3 Teorema

Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ es diferenciable en $\mathbf{x}_0$ y $\mathcal{R}^n \xrightarrow{g} \mathcal{R}^n$ es diferenciable en $\mathbf{y}_0 = f(\mathbf{x}_0)$, entonces el determinante jacobiano de $g \circ f$ en $\mathbf{x}_0$ es el producto del determinante jacobiano de f en $\mathbf{x}_0$ y el de g en $\mathbf{y}_0$.

Si f está definida por

$$f\begin{pmatrix}x_1\\ \cdot\\ \cdot\\ \cdot\\ x_n\end{pmatrix} = \begin{pmatrix}f_1(x_1, \ldots, x_n)\\ \cdot\\ \cdot\\ \cdot\\ f_n(x_1, \ldots, \ _n)\end{pmatrix} = \begin{pmatrix}y_1\\ \cdot\\ \cdot\\ \cdot\\ y_n\end{pmatrix},$$

entonces el determinante jacobiano $\det f'$ se denota a menudo por

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)},$$

o, equivalentemente, por

$$\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}.$$

Ejemplo 7. Sean

$$f\begin{pmatrix}r\\ \theta\end{pmatrix} = \begin{pmatrix}r\cos\theta\\ r\,\mathrm{sen}\,\theta\end{pmatrix} = \begin{pmatrix}x\\ y\end{pmatrix} \quad \text{y} \quad g\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}x^2 - y^2\\ 2xy\end{pmatrix} = \begin{pmatrix}w\\ z\end{pmatrix}.$$

Entonces

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det\begin{pmatrix}\cos\theta & -r\,\mathrm{sen}\,\theta\\ \mathrm{sen}\,\theta & r\cos\theta\end{pmatrix} = r(\cos^2\theta + \mathrm{sen}^2\theta) = r,$$

y

$$\frac{\partial(w, z)}{\partial(x, y)} = \det\begin{pmatrix}2x & -2y\\ 2y & 2x\end{pmatrix} = 4(x^2 + y^2).$$

El determinante jacobiano de la función compuesta $g \circ f$ se denota en este caso por $\partial(w, z)/\partial(r, \theta)$. Si

$$\begin{pmatrix}x_0\\ y_0\end{pmatrix} = \begin{pmatrix}r_0\cos\theta_0\\ r_0\,\mathrm{sen}\,\theta_0\end{pmatrix},$$

entonces el Teorema 3.3 implica que

$$\begin{aligned}\frac{\partial(w, z)}{\partial(r, \theta)}(r_0, \theta_0) &= \frac{\partial(w, z)}{\partial(x, y)}(x_0, y_0)\frac{\partial(x, y)}{\partial(r, \theta)}(r_0, \theta_0)\\ &= 4(x_0^2 + y_0^2)r_0 = 4r_0^3.\end{aligned}$$

EJERCICIOS

1. Si

$$f\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}x^2 + xy + 1\\ y^2 + 2\end{pmatrix}, \quad g\begin{pmatrix}u\\ v\end{pmatrix} = \begin{pmatrix}u + v\\ 2u\\ v^2\end{pmatrix},$$

encuentre la matriz de la diferencial de la función compuesta $g \circ f$ en

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \qquad \left[Resp: \begin{pmatrix} 3 & 3 \\ 6 & 2 \\ 0 & 12 \end{pmatrix}. \right]$$

2. Sean

$$f(t) = \begin{pmatrix} t \\ t+1 \\ t^2 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y

$$g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z^2 \\ x^2 - y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}.$$

(a) Encuentre la matriz jacobiana de $g \circ f$ en $t = a$. $\left[Resp: \begin{pmatrix} 3 + 4a^3 \\ 2a - 1 \end{pmatrix}. \right]$

(b) Calcule du/dt en términos de las derivadas de x, y, z y las derivadas parciales de u.

3. Considere la curva definida paramétricamente por

$$f(t) = \begin{pmatrix} t \\ t^2 - 4 \\ e^{t-2} \end{pmatrix}, \qquad -\infty < t < \infty.$$

Sea g una función real, diferenciable y con dominio $\mathbb{R}^3$. Si

$$\mathbf{x}_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

y

$$\frac{\partial g}{\partial x}(\mathbf{x}_0) = 4, \qquad \frac{\partial g}{\partial y}(\mathbf{x}_0) = 2, \qquad \frac{\partial g}{\partial z}(\mathbf{x}_0) = 2,$$

[*Resp.* 14.]

calcule $d(g \circ f)/dt$ en $t = 2$.

4. Considere las funciones

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ u - v \\ u^2 - v^2 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y

$$F(x, y, z) = x^2 + y^2 + z^2 = w.$$

(a) Calcule la matriz que define la diferencial de $F \circ f$ en $\begin{pmatrix} a \\ b \end{pmatrix}$.

(b) Calcule $\partial w/\partial u$ y $\partial w/\partial v$.

5. Sea $u = f(x, y)$. Haga los cambios de variables $x = r \cos\theta$, $y = r \operatorname{sen}\theta$. Si sabemos que

$$\frac{\partial f}{\partial x} = x^2 + 2xy - y^2 \quad \text{y} \quad \frac{\partial f}{\partial y} = x^2 - 2xy + 2,$$

[*Resp.* 8.]

calcule $\partial f/\partial\theta$, cuando $r = 2$ y $\theta = \pi/2$.

6. Si $w = \sqrt{x^2 + y^2 + z^2}$ y

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\operatorname{sen}\theta \\ r \end{pmatrix},$$

encuentre $\partial w/\partial r$ y $\partial w/\partial\theta$ usando la regla de la cadena. Compruebe su resultado efectuando las sustituciones.

7. las funciones vectoriales f y g están definidas por

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u\cos v \\ u \operatorname{sen} v \end{pmatrix}, \qquad \begin{matrix} 0 < u < \infty, \\ -\pi/2 < v < \pi/2, \end{matrix}$$

$$g\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \operatorname{arctg} \dfrac{y}{x} \end{pmatrix}, \qquad 0 < x < \infty.$$

(a) Calcule la matriz jacobiana de $g \circ f$ en $\begin{pmatrix} u \\ v \end{pmatrix}$.

(b) Calcule la matriz jacobiana de $f \circ g$ en $\begin{pmatrix} x \\ y \end{pmatrix}$.

(c) ¿Son verdaderas o falsas las siguientes afirmaciones?
1. dominio de f = dominio de $g \circ f$.
2. dominio de g = dominio de $f \circ g$.

8. Una función I se llama una **función identidad** si $I(\mathbf{x}) = \mathbf{x}$ para todo $\mathbf{x}$ en el dominio de I.

(a) Demuestre que si las funciones diferenciables f y g son tales que la compuesta $g \circ f$ es una función identidad, entonces la transformación $(d_{f(\mathbf{x})}g) \circ (d_{\mathbf{x}}f)$ es también una función identidad para todo $\mathbf{x}$ en el dominio de $g \circ f$.

(b) ¿Qué relación tiene este ejercicio con el anterior?

9. Sea $\mathbf{x}_1$ un vector tangente en $\mathbf{x}_0$ a la curva definida paramétricamente por la función vectorial diferenciable f. Si $\mathbf{x}_0$ está en el dominio de la función vectorial diferenciable F, demuestre que si $F'(\mathbf{x}_0)\mathbf{x}_1$ no es cero, entonces él es un vector tangente en $F(\mathbf{x}_0)$ a la curva definida paramétricamente por $F \circ f$.

10. La convención de denotar las funciones coordenadas por variables reales tiene sus trampas. Aclare la paradoja siguiente: Sea $w = f(x, y, z)$ y $z = g(x, y)$. Por la regla de la cadena,

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x}.$$

x e y no están relacionadas, por tanto, $\partial y/\partial x = 0$. Además es claro que $\partial x/\partial x = 1$ y, por tanto,

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x},$$

de modo que

$$0 = \frac{\partial w}{\partial z}\frac{\partial z}{\partial x}.$$

En particular, tomemos $w = 2x + y + 32$ y $t = y + 18$. Por tanto,

$$\frac{\partial w}{\partial z} = 3 \quad \text{y} \quad \frac{\partial z}{\partial x} = 5.$$

y entonces $0 = 15$, lo cual, evidentemente, es falso.

11. Si $y = f(x - at) + g(x + at)$, donde a es una constante y f y g son dos veces diferenciables, demuestre que

$$a^2\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}. \qquad \text{(Ecuación de onda)}$$

12. Si $z = f(x, y)$ es diferenciable y $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\text{sen}\,\theta \end{pmatrix}$, demuestre que

$$\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2.$$

13. Si $f(tx, ty) = t^n f(x, y)$ para algún entero n y todo x, y, t, demuestre que

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = nf(x, y).$$

14. Demuestre que si $g(x, y)$ es una función diferenciable, entonces

$$\frac{dg(x, x)}{dx} = \frac{\partial g}{\partial x}(x, x) + \frac{\partial g}{\partial y}(x, x).$$

Demuestre, usando la función $f(x) = (x, x)$, que este resultado es equivalente a $(g \circ f)'(x) = g'(f(x))f'(x)$.
Aplique la ecuación de la función $g(x, y) = x^y$.

15. (a) Si

$$w = f(x, y, z, t), \qquad x = g(u, z, t), \qquad \text{y} \qquad z = h(u, t),$$

encuentre una fórmula para dw/dt, símbolo que significa la velocidad de cambio de

w con respecto a t, tomando en consideración todas las interrelaciones entre w, x, z, t.
(b) Si

$$f(x, y, z, t) = 2xy + 3z + t^2,$$
$$g(u, z, t) = ut\,\text{sen}\,z,$$
$$h(u, t) = 2u + t,$$

evalúe dw/dt en el punto $u = 1$, $t = 2$, $y = 3$, usando la fórmula que dedujo en (a) y, también, sustituyendo en las funciones para x y z y luego derivando.

16. Considere una función real $f(x, y)$ tal que

$$f_x(2, 1) = 3, \quad f_y(2, 1) = -2, \quad f_{xx}(2, 1) = 0,$$
$$f_{xy}(2, 1) = f_{yx}(2, 1) = 1, \quad f_{yy}(2, 1) = 2.$$

Sea $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$ definida por

$$g(u, v) = (u + v, uv).$$

Encuentre $\partial^2(f \circ g)/\partial v\,\partial u$ en $(1, 1)$. [*Resp.* 2.]

17. Calcule los determinantes jacobianos de las siguientes funciones en los puntos indicados.

(a) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 + 2uv + 3v \\ u - v \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$, a $\mathbf{u}_0 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$. [*Resp.* -7.]

(b) $g\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = \begin{pmatrix} z \\ w \end{pmatrix}$, a $\mathbf{x}_0 = \begin{pmatrix} 6 \\ -2 \end{pmatrix}$. [*Resp.* 160.]

(c) $A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$, en un $\begin{pmatrix} x \\ y \end{pmatrix}$ arbitrario.

(d) Una transformación afín $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$, $A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0$ en un $\mathbf{x}_0$ arbitrario.

(e) $T\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta\,\text{sen}\,\phi \\ r\,\text{sen}\,\theta\,\text{sen}\,\phi \\ r\cos\phi \end{pmatrix}$, a $\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix}$.

18. Si f y g son las funciones del Ejercicio 17(a) y (b), calcule el determinante jacobiano de la función compuesta $g \circ f$ en $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

[*Resp.* -1120.]

19. ¿Qué dicen las ecuaciones en términos de $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2$ y $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$ y cómo se obtienen a partir del Teorema 3.3?

(a) $\dfrac{\partial(x, y)}{\partial(u, v)}\dfrac{\partial(u, v)}{\partial(r, \theta)} = \dfrac{\partial(x, y)}{\partial(r, \theta)}$. (b) $\dfrac{\partial(x, y)}{\partial(u, v)}\dfrac{\partial(u, v)}{\partial(x, y)} = 1$.

4. Funciones definidas implícitamente

Es posible que los vectores estén relacionados por una fórmula que no exprese directamente a uno como función del otro. Por ejemplo, la fórmula

$$\frac{pv}{t} = k_0$$

expresa la relación entre la presión p y el volumen y la temperatura (v, t) de un gas en un recipiente. También, las ecuaciones

$$x^2 + y^2 + z^2 = 1$$
$$x + y + z = 0$$

pueden ser interpretadas como una relación entre las tres coordenadas de un punto en un círculo de radio 1 con centro $(0, 0, 0)$ en $\mathcal{R}^3$. En estos ejemplos, las ecuaciones no dan una fórmula explícita para una de las coordenadas en términos de la otra. En esta sección consideraremos la aplicación del cálculo a este tipo de relaciones.

Dadas dos funciones $\mathcal{R}^2 \xrightarrow{F} \mathcal{R}$ y $\mathcal{R} \xrightarrow{f} \mathcal{R}$, **la ecuación**

$$F(x, y) = 0 \tag{1}$$

define f **implícitamente** si $F(x, f(x)) = 0$ para todo x en el dominio de f. El cero de la Ec. (1) puede ser remplazado por cualquier constante c. Sin embargo, como $F(x, y) = c$ es equivalente a $G(x, y) = F(x, y) - c = 0$, es habitual incluir la constante c en la función F.

Ejemplo 1. Sea $F(x, y) = x^2 + y^2 - 1$. Entonces la condición $F(x, f(x)) = x^2 + (f(x))^2 - 1 = 0$ para todo x en el dominio de f es satisfecha por las siguientes funciones

$$f_1(x) = \sqrt{1 - x^2}, \qquad -1 \le x \le 1.$$
$$f_2(x) = -\sqrt{1 - x^2}, \qquad -1 \le x \le 1.$$
$$f_3(x) = \begin{cases} \sqrt{1 - x^2}, & -\tfrac{1}{2} \le x < 0. \\ -\sqrt{1 - x^2}, & 0 \le x \le 1. \end{cases}$$

La Fig. 7 muestra sus gráficos. De la definición de una función definida implícitamente se deduce que f_1, f_2, f_3 están definidas implícitamente por la ecuación $x^2 + y^2 - 1 = 0$.

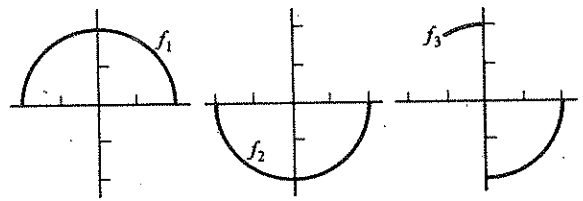


Figura 7

Consideremos la función $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$. Un elemento arbitrario de $\mathcal{R}^{n+m}$ puede ser expresado como $(x_1, \ldots, x_n, y_1, \ldots, y_m)$ o como un par $(\mathbf{x}, \mathbf{y})$, donde $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_m)$. De esta manera, F puede ser considerada como una función de dos variables vectoriales, $\mathbf{x}$ en $\mathcal{R}^n$ e $\mathbf{y}$ en $\mathcal{R}^m$ o, como una función de un vector variable $(\mathbf{x}, \mathbf{y})$ en $\mathcal{R}^{n+m}$. La función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ está definida implícitamente por la ecuación

$$F(\mathbf{x}, \mathbf{y}) = 0$$

si $F(\mathbf{x}, f(\mathbf{x})) = 0$ para todo $\mathbf{x}$ en el dominio de f.

Ejemplo 2. Las ecuaciones

$$\begin{aligned} x + y + z - 1 &= 0 \\ 2x \quad + z + 2 &= 0 \end{aligned} \tag{2}$$

determinan y y z como funciones de x. Obtenemos

$$y = x + 3, \qquad z = -2x - 2.$$

En términos de una función $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}^2$, las Ecs. (2) tienen la forma

$$F\left(x, \begin{pmatrix} y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + y + z - 1 \\ 2x \quad + z + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} x + \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

La función $\mathcal{R} \xrightarrow{f} \mathcal{R}^2$ definida implícitamente es

$$f(x) = \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} x + 3 \\ -2x - 2 \end{pmatrix}.$$

A pesar de que el Ejemplo 1 muestra que una función definida implícitamente no es necesariamente continua, en esta sección nos preocuparemos principalmente de funciones que no solo son continuas, sino también diferenciables. El teorema de la función implícita, que se dará al final de la Sec. 6, da condiciones para la existencia de una función diferenciable f definida por la ecuación $F(\mathbf{x}, f(\mathbf{x})) = 0$. Antes de discutir ese teorema, consideraremos el problema de encontrar la diferencial de f si se sabe que ella existe. Supongamos que la función $\mathcal{R}^2 \xrightarrow{F} \mathcal{R}$ y $\mathcal{R} \xrightarrow{f} \mathcal{R}$ son diferenciables y que

$$F(x, f(x)) = 0$$

para todo x en el dominio de f. Entonces la regla de la cadena aplicada a $F(x, f(x))$, se tiene

$$F_x(x, f(x)) + F_y(x, f(x))f'(x) = 0.$$

Por tanto,

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}, \qquad \text{si} \quad F_y(x, f(x)) \neq 0.$$

Para funciones vectoriales es posible hacer un cálculo similar.

Ejemplo 3. Dadas las ecuaciones

$$x^2 + y^2 + z^2 - 5 = 0, \qquad xyz + 2 = 0, \tag{3}$$

supongamos que x e y son funciones diferenciables de z, es decir, que la función definida implícitamente por la Ec. (3) es de la forma $(x, y) = f(z)$. Para calcular dx/dz y dy/dz, aplicamos la regla de la cadena a las ecuaciones dadas, tenemos que

$$2x\frac{dx}{dz} + 2y\frac{dy}{dz} + 2z = 0,$$

$$yz\frac{dx}{dz} + xz\frac{dy}{dz} + xy = 0.$$

Estas nuevas ecuaciones pueden ser resueltas, obteniendo la solución

$$\begin{pmatrix} \frac{dx}{dz} \\ \frac{dy}{dz} \end{pmatrix} = \begin{pmatrix} \frac{x(y^2 - z^2)}{z(x^2 - y^2)} \\ \frac{y(z^2 - x^2)}{z(x^2 - y^2)} \end{pmatrix},$$

que es la matriz de $f'(z)$. Nótese que los valores que corresponden a x e y deben ser conocidos para hacer explícita esta fórmula. Es decir, no es posible evaluar $(dx/dz)(1)$ a partir de la información dada. Por otra parte, dado el punto $(x, y, z) = (1, -2, 1)$, tenemos que $(dx/dz)(1) = -1$. La razón para esto es que, tal como en el Ejemplo 1, hay más de una función f definida implícitamente por las Ecs. (3). Si especificamos un punto particular en su gráfico, f queda determinada unívocamente en una vecindad de este punto.

Ejemplo 4. Consideremos

$$xu + yv + zw = 1,$$

$$x + y + z + u + v + w = 0,$$

$$xy + zuv + w = 1.$$

Supongamos que x, y y z son funciones de u, v y w. Para encontrar las derivadas de x, y y z con respecto a w, derivamos las tres ecuaciones usando la regla de la cadena

$$u\frac{\partial x}{\partial w} + v\frac{\partial y}{\partial w} + w\frac{\partial z}{\partial w} + z = 0,$$

$$\frac{\partial x}{\partial w} + \frac{\partial y}{\partial w} + \frac{\partial z}{\partial w} + 1 = 0,$$

$$y\frac{\partial x}{\partial w} = x\frac{\partial y}{\partial w} + uv\frac{\partial z}{\partial w} + 1 + 0.$$

De donde

$$\frac{\partial x}{\partial w} = \frac{uv^2 + xz + w - zuv - xw - v}{u^2v + vy + wx - yw - ux - uv^2}.$$

En forma análoga, podemos encontrar $\partial y/\partial w$ y $\partial z/\partial w$. Para encontrar la derivada parcial con respecto a u, derivamos las ecuaciones originales con respecto a u y despejamos $\partial x/\partial u$, $\partial y/\partial u$ y $\partial z/\partial u$. Las derivadas parciales con respecto a v se encuentran de la misma forma.

Los cálculos hechos en el Ejemplo 4 proporcionan los nueve elementos de la matriz de la diferencial de una función vectorial definida implícitamente. Para que estos cálculos funcionen, es necesario tener un número de ecuaciones igual al número de funciones coordenadas definidas implícitamente. Para ver mejor las razones para este requisito, supongamos dada una función vectorial diferenciable

$$F(u, v, x, y) = \begin{pmatrix} F_1(u, v, x, y) \\ F_2(u, v, x, y) \end{pmatrix}$$

y que las ecuaciones

$$F_1(u, v, x, y) = 0, \qquad F_2(u, v, x, y) = 0 \tag{4}$$

definen implícitamente una función diferenciable $(x, y) = f(u, v)$. Si derivamos las Ecs. (4) con respecto de u y v, usando la regla de la cadena, obtenemos

$$\frac{\partial F_1}{\partial u} + \frac{\partial F_1}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial F_1}{\partial y}\frac{\partial y}{\partial u} = 0, \qquad \frac{\partial F_1}{\partial v} + \frac{\partial F_1}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial F_1}{\partial y}\frac{\partial y}{\partial v} = 0,$$

$$\frac{\partial F_2}{\partial u} + \frac{\partial F_2}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial F_2}{\partial y}\frac{\partial y}{\partial u} = 0, \qquad \frac{\partial F_2}{\partial v} + \frac{\partial F_2}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial F_2}{\partial y}\frac{\partial y}{\partial v} = 0.$$

Estas ecuaciones pueden ser escritas en forma matricial como sigue:

$$\begin{pmatrix} \frac{\partial F_1}{\partial u} & \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial u} & \frac{\partial F_2}{\partial v} \end{pmatrix} + \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = 0. \tag{5}$$

La última matriz de la derecha es la matriz de la diferencial de f en (u, v). Si la despejamos, tenemos

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = - \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial u} & \frac{\partial F_2}{\partial v} \end{pmatrix}. \tag{6}$$

Para poder resolver en forma única la Ec. (5) en $f'(u, v)$, es esencial que exista la matriz inversa que aparece en la Ec. (6). En particular, esto implica que el número de ecuaciones dadas originalmente sea igual al número de variables determinadas implícitamente o, lo que es lo mismo, que los espacios imagen de F y f tengan la misma dimensión.

Una ecuación análoga a la Ec. (6) es válida para un número arbitrario de funciones coordenadas F_i y se demuestra exactamente de la misma manera. Podemos resumir este resultado en la siguiente generalización de 4.1.

4.2 Teorema

Si $\mathbb{R}^{n+m} \xrightarrow{F} \mathbb{R}^m$ y $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$ son funciones diferenciables e $\mathbf{y} = f(\mathbf{x})$ satisface $F(\mathbf{x}, \mathbf{y}) = 0$, entonces

$$f'(\mathbf{x}) = -F_{\mathbf{y}}^{-1}(\mathbf{x}, f(\mathbf{x}))F_{\mathbf{x}}(\mathbf{x}, f(\mathbf{x})),$$

siempre que $F_{\mathbf{y}}$ tenga inverso. La derivada $F_{\mathbf{y}}$ se calcula manteniendo $\mathbf{x}$ fijo y $F_{\mathbf{x}}$ manteniendo $\mathbf{y}$ fijo.

La notación usada se ilustra en el ejemplo siguiente.

Ejemplo 5. Supongamos que

$$F(x, y, z) = \begin{pmatrix} x^2y + xz \\ xz + yz \end{pmatrix}$$

y que elegimos $\mathbf{x} = x$, $\mathbf{y} = (y, z)$. Entonces

$$F_x(x, y, z) = \begin{pmatrix} 2xy + z \\ z \end{pmatrix}$$

y

$$F_{(y,z)}(x, y, z) = \begin{pmatrix} x^2 & x \\ z & x + y \end{pmatrix}.$$

El método de Newton es, en muchos casos, una manera efectiva para encontrar un valor aproximado de una función definida implícitamente. Supongamos que $f(\mathbf{x})$ satisface la ecuación $F(x, f(\mathbf{x})) = 0$ en una vecindad de $\mathbf{x} = \mathbf{x}_0$. Para calcular $f(\mathbf{x}_0)$ debemos resolver la ecuación $F(\mathbf{x}_0, \mathbf{y}) = 0$ para $\mathbf{y}$. Aplicamos el método de Newton a la función $g(\mathbf{y}) = F(\mathbf{x}_0, \mathbf{y})$. La ecuación de iteración $\mathbf{y}_{k+1} = \mathbf{y}_k - [g'(\mathbf{y}_k)]^{-1}g(\mathbf{y}_k)$ toma la forma

4.3 $$\mathbf{y}_{k+1} = \mathbf{y}_k - [F_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_k)]^{-1}F(\mathbf{x}_0, \mathbf{y}_k).$$

Por supuesto que hay que hacer una elección de $\mathbf{y}_0$ usando información más detallada acerca de F.

Ejemplo 6. Si $F(x, y) = x^2y + xy^2$, entonces $F_y(x, y) = x^2 + 2xy$ y la Ec. 4.3 toma la forma

$$y_{k+1} = y_k - \frac{x_0^2y_k + x_0y_k^2}{x_0^2 + 2x_0y_k} = \frac{y_k^2}{x_0 + 2y_k}. \tag{7}$$

Supongamos que $x_0 = 1$ y que elegimos $y_0 = 1$ como aproximación inicial del mismo $f(x_0)$ que satisface $F(x_0, f(x_0)) = 0$. Entonces

$$y_1 = \tfrac{1}{3},$$

$$y_2 = \frac{(\frac{1}{3})^2}{1 + 2(\frac{1}{3})} = \frac{1}{15},$$

$$y_3 = \frac{(\frac{1}{15})^2}{1 + 2(\frac{1}{15})} = \frac{1}{255}.$$

Estos valores de y sugieren que $y = 0$ puede ser una solución y se puede verificar que, en efecto, $F(1, 0) = 0$.

Por otra parte, si probamos con $y_0 = -2$ y todavía $x_0 = 1$, obtenemos

$$y_1 = \frac{(-2)^2}{1 + 2(-2)} = -\frac{4}{3},$$

$$y_2 = \frac{(-\frac{4}{3})^2}{1 + 2(-\frac{4}{3})} = -\frac{16}{15},$$

$$y_3 = \frac{(-\frac{16}{15})^2}{1 + 2(-\frac{16}{15})} = -\frac{256}{255}.$$

Estos valores de y parecen converger a -1 y, en efecto, se tiene que $F(1, -1) = 0$. Para resolver aproximadamente $F(1{,}1, y)$ tomamos $x_0 = 1{,}1$ en la Ec. (7) y probamos con y_0 o $y_0 = 1$.

Ejemplo 7. Supongamos que

$$F(x, y, z) = \begin{pmatrix} x^2y + z \\ x + y^2z \end{pmatrix}$$

y que queremos encontrar valor aproximado de y y z tales que $F(1, y, z) = 0$. La matriz derivada de F con respecto de (y, z) es

$$F_{(y,z)}(x, y, z) = \begin{pmatrix} x^2 & 1 \\ 2yz & y^2 \end{pmatrix};$$

de modo que

$$F_{(y,z)}(1, y, z) = \begin{pmatrix} 1 & 1 \\ 2yz & y^2 \end{pmatrix}.$$

Si calculamos la matriz inversa mediante la fórmula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

obtenemos

$$[F_{(y,z)}(1, y, z)]^{-1} = \frac{1}{y^2 - 2yz} \begin{pmatrix} y^2 & -1 \\ -2yz & 1 \end{pmatrix}.$$

Entonces la fórmula de iteración 4.3 da

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} y_k \\ z_k \end{pmatrix} - \frac{1}{y_k^2 - 2y_kz_k} \begin{pmatrix} y_k^2 & -1 \\ -2y_kz_k & 1 \end{pmatrix} \begin{pmatrix} y_k + z_k \\ 1 + y_k^2z_k \end{pmatrix}$$

$$= \frac{1}{y_k^2 - 2y_kz_k} \begin{pmatrix} 1 - 2y_k^2z_k \\ -1 + 2y_k^2z_k \end{pmatrix}.$$

Esta fórmula muestra que si elegimos arbitrariamente y_0 y z_0 de modo que $y_0^2 - 2y_0z_0 \neq 0$, tenemos que $y_{k+1} = -z_{k+1}$ para $k \geq 0$. Por tanto, podemos probar con $y_0 = 1$, $z_0 = -1$. Si sustituimos en la fórmula de iteración, obtenemos que y_{k+1}, $z_{k+1} = -1$ para todo $k \geq 0$ y tenemos que en realidad $F(0, 1, 1) = 0$. Por tanto, en este ejemplo no es necesario hacer otros cálculos numéricos debido a la elección afortunada de y_0 y z_0. Si no hubiéramos sido tan afortunados, habríamos tenido que calcular aproximaciones sucesivas (x_k, y_k) de la solución.

EJERCICIOS

1. Si

$$x^2y + yz = 0, \qquad xyz + 1 = 0,$$

encuentre dx/dz y dy/dz en $(x, y, z) = (1, 1, -1)$.

$$\left[\textit{Resp: } \frac{dx}{dz} = -\frac{1}{2}, \frac{dy}{dz} = \frac{3}{2}. \right]$$

2. Si expresamos el Ejercicio 1 en la notación vectorial del Teorema 4.2, ¿cuáles son F, x, y, F_y y F_x?

3. Si

$$x + y - u - v = 0,$$
$$x - y + 2u + v = 0,$$

calcule $\partial x/\partial u$ y $\partial y/\partial u$:

(a) Expresando x e y en términos de u y v.
(b) Derivando implícitamente.

4. Si expresamos el Ejercicio 3 en la notación vectorial del Teorema 4.2, ¿cuál es la matriz de $f'(\mathbf{x})$?

5. Si $x^2 + yu + xv + w = 0$, $x + y + uvw + 1 = 0$, entonces, considerando x e y como funciones de u, v y w, encuentre

$$\frac{\partial x}{\partial u} \quad \text{y} \quad \frac{\partial y}{\partial u} \quad \text{a} \quad (x, y, u, v, w) = (1, -1, 1, 1, -1).$$

$$\left[\textit{Resp: } \frac{\partial x}{\partial u} = 0, \frac{\partial y}{\partial u} = 1. \right]$$

6. (a) Las ecuaciones $2x^3y + yx^2 + t^2 = 0$, $x + y + t - 1 = 0$ definen implícitamente una curva

$$f(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

que satisface

$$f(1) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Encuentre la recta tangente a f en $t = 1$.

$$\left[\textit{Resp: } t\begin{pmatrix} -\frac{3}{5} \\ -\frac{2}{5} \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \right]$$

(b) Use el método de Newton para aproximar $f(1, 1)$.

7. Si la ecuación $x^2/4 + y^2 + z^2/9 - 1 = 0$ define z implícitamente como función $z = f(x, y)$ cerca del punto $x = 1$, $y = \sqrt{11/6}$, $z = 2$. El gráfico de la función f es

una superficie. Encuentre el plano tangente en $(1, \sqrt{11/6}, 2)$.

8. Supongamos que la ecuación $F(x, y, z) = 0$ define implícitamente $z = f(x, y)$ y que $z_0 = f(x_0, y_0)$. Supongamos, además, que la superficie que es el gráfico de $z = f(x, y)$ tiene un plano tangente en (x_0, y_0). Demuestre que

$$(x - x_0)\frac{\partial F}{\partial x}(x_0, y_0, z_0) + (y - y_0)\frac{\partial F}{\partial y}(x_0, y_0, z_0) + (z - z_0)\frac{\partial F}{\partial z}(x_0, y_0, z_0) = 0$$

es la ecuación de este plano tangente.

9. Las ecuaciones

$$\begin{aligned} 2x + y + 2z + y - v - 1 &= 0 \\ xy + z - u + 2v - 1 &= 0 \\ yz + xz + u^2 + v &= 0 \end{aligned}$$

definen x, y y z como funciones de u y v cerca de $(x, y, z, u, v) = (1, 1, -1, 1)$.

(a) Encuentre la matriz de la diferencial de la función definida implícitamente

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix} = f(u, v) \quad \text{a} \quad (u, v) = (1, 1).$$

(b) La función f define paramétricamente una superficie en el espacio (x, y, z). Encuentre el plano tangente a ella en el punto $(1, 1, -1)$.

(c) Aplique el método de Newton para aproximar $f(1{,}1, 1{,}1)$.

5. Coordenadas curvilíneas

A menudo, las fórmulas que aparecen en las matemáticas y sus aplicaciones pueden ser simplificadas si se hace una descripción adecuada de las magnitudes que se deben considerar especialmente. Como generalmente, en la práctica, estas magnitudes están expresadas mediante vectores con coordenadas reales, el problema puede ser considerado como el de elegir el sistema de coordenadas más adecuado. Por tanto, consideramos la introducción de coordenadas en $\mathcal{R}^n$ distintas de las coordenadas naturales x_i que aparecen en la designación de un punto típico $(x_1, \ldots, x_n)$. En forma precisa, a cada punto $(x_1, \ldots, x_n)$ se le asociará una nueva n-pla $(u_1, \ldots, u_n)$. Es claro que si queremos pasar de uno a otro de estos conjuntos de coordenadas, es necesario que la asignación recién descrita sea biunívoca, es decir, para cada $(x_1, \ldots, x_n)$ debe existir un solo $(u_1, \ldots, u_n)$ y viceversa. En la práctica, la asignación de nuevas coordenadas se hace a menudo en una subregión específica de $\mathcal{R}^n$ y no en todo el espacio. El espacio vectorial de las nuevas coordenadas $(u_1, \ldots, u_n)$ será denotado por

$\mathcal{U}^n$ para no confundir con $\mathcal{R}^n$, a cuyos puntos $(x_1, \ldots, x_n)$ se le están asignando nuevas coordenadas.

Ejemplo 1. (*Coordenadas polares en el plano.*) Consideremos dos copias del espacio de dos dimensiones: el plano xy, denotado por $\mathcal{R}^2$ y el plano $r\theta$ denotado por $\mathcal{U}^2$. La función $\mathcal{U}^2 \xrightarrow{T} \mathcal{R}^2$ definida por

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\mathrm{sen}\,\theta \end{pmatrix}, \qquad \begin{cases} 0 < r < \infty \\ -\infty < \theta < \infty \end{cases} \tag{1}$$

tiene una descripción geométrica simple. La imagen por T de un punto $\begin{pmatrix} r \\ \theta \end{pmatrix}$ es el punto

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

cuya distancia al origen es r y cuyo ángulo con el eje x positivo en la dirección de los punteros del reloj es θ. Véase la Fig. 8.

Dados dos puntos

$$\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}$$

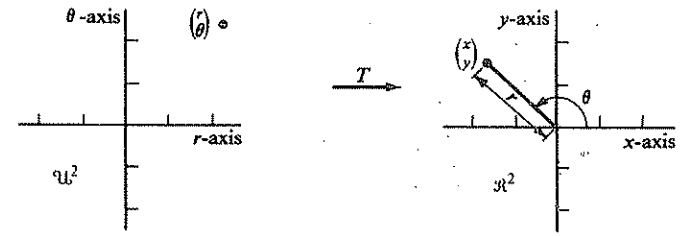


Figura 8

en el dominio de T, es fácil demostrar que

$$T\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}$$

si, y solo si, $r_1 = r_2$ y $\theta_1 = \theta_2 + 2\pi m$ para algún entero m. La imagen de T está formada por todos $\mathcal{R}^2$, salvo el origen. Por tanto, para todo punto

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{en } \mathcal{R}^2 \text{ excepto } \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

existen números r y θ llamadas **coordenadas polares** de $\mathbf{x}$, tales que

$$T\begin{pmatrix} r \\ \theta \end{pmatrix} = \mathbf{x}.$$

Además, las coordenadas polares de $\mathbf{x}$ están especificadas unívocamente, salvo un múltiplo de 2π en la segunda coordenada.

Desde un punto de vista ligeramente diferente, el párrafo anterior dice que T no es inyectiva, pero que sí lo es si su dominio se restringe a un subconjunto del conjunto en el plano $r\theta$ definido por

$$0 < r < \infty, \qquad \theta_0 \leq \theta < \theta_0 + 2\pi.$$

Cuando T se restringe de esta manera, tiene una función inversa. Si resolvemos las ecuaciones $x = r\cos\theta$ y $y = r\,\text{sen}\,\theta$, si $x \neq 0$, obtenemos

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \text{arctg}\,\frac{y}{x} + k\pi.$$

Hemos usado la convención habitual de que la inversa de una función trigonométrica es la rama principal de la función multivalente correspondiente. Por tanto, la imagen de la función arctg es el intervalo $-\pi/2 < \theta < \pi/2$. De esto sigue que la función definida por

$$\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \text{arctg}\,\dfrac{y}{x} \end{pmatrix}, \qquad x > 0,$$

es la inversa de la restricción de T a la región $0 < r < \infty$, $-\pi/2 < \theta < \pi/2$. En forma análoga, la función definida por

$$\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \text{arcotg}\,\dfrac{x}{y} \end{pmatrix}, \qquad y > 0,$$

es la inversa de la restricción de T a $0 < r < \infty$, $0 < \theta < \pi$.

No hemos definido las coordenadas polares del origen del plano xy por la sencilla razón que

$$\begin{pmatrix} 0\cos\theta \\ 0\,\text{sen}\,\theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ para todo } \theta,$$

y, por tanto, el requisito de inyectividad falla en el origen. Este hecho no causa ninguna dificultad real. Por ejemplo, la ecuación de la lemniscata en coordenadas rectangulares

$$(x^2 + y^2)^2 = 2(x^2 - y^2), \tag{2}$$

se transforma, al tomar coordenadas polares, en

$$r^2 = 2\cos 2\theta, \qquad r > 0. \tag{3}$$

La imagen por T del conjunto de los pares (r, θ) que satisfacen la Ec. (3) es precisamente el conjunto de los pares (x, y) distintos del origen que satisfacen la Ec. (2). Podemos simplemente agregar este punto. Véase la Fig. 9.

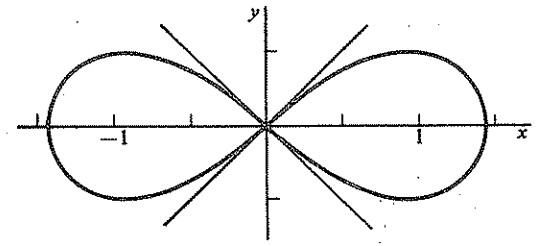


Figura 9

Ejemplo 2. (*Coordenadas esféricas en el espacio de* 3 *dimensiones.*) Consideremos la función $\mathfrak{U}^3 \xrightarrow{T} \mathfrak{R}^3$ definida por

$$T\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} r\,\text{sen}\,\phi\cos\theta \\ r\,\text{sen}\,\phi\,\text{sen}\,\theta \\ r\cos\phi \end{pmatrix}, \qquad \begin{cases} 0 < r < \infty \\ 0 < \phi < \pi \\ 0 \leq \theta < 2\pi. \end{cases} \tag{4}$$

Por mayor simplicidad, hemos restringido desde un principio el dominio de T de modo que ésta sea inyectiva. Su imagen es todo $\mathfrak{R}^3$ con la excepción del eje z. Por tanto, se asignan **coordenadas esféricas** (r, ϕ, θ) a todos los puntos de $\mathfrak{R}^3$ que no están en el eje z. Como en el caso de las coordenadas polares, las coordenadas esféricas (r, ϕ, θ) de un punto $\mathbf{x} = (x, y, z)$ tienen una interpretación geométrica sencilla (véase Fig. 10): El número r es la distancia de

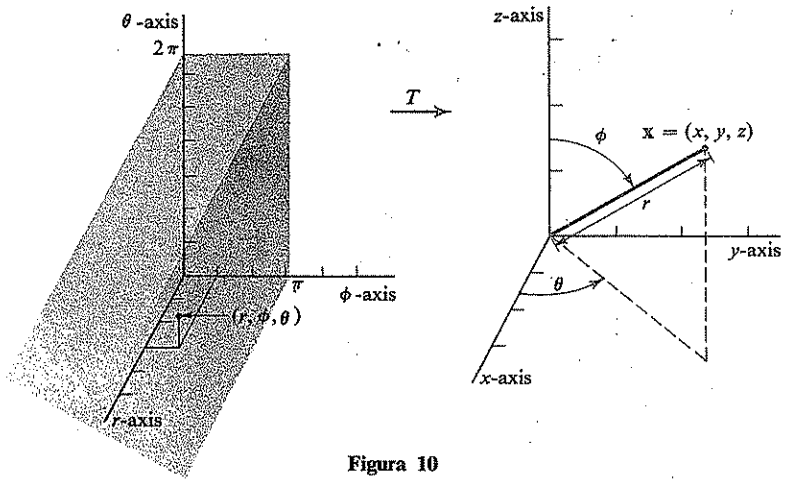


Figura 10

x al origen. La coordenada ϕ es el ángulo en radianes entre el vector x y el eje z positivo. Por último, θ es el ángulo en radianes del eje x positivo a la proyección $(x, y, 0)$ de x en el plano xy.

Podemos obtener una expresión explícita para la función inversa, que denotaremos por T^{-1}, resolviendo en r, θ y ϕ las ecuaciones

$$\begin{aligned} x &= r\,\text{sen}\,\phi\cos\theta, \\ y &= r\,\text{sen}\,\phi\,\text{sen}\,\theta, \\ z &= r\cos\phi, \end{aligned}$$

Obtenemos para $y \geq 0$,

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = T^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sqrt{x^2+y^2+z^2} \\ \arccos\dfrac{z}{\sqrt{x^2+y^2+z^2}} \\ \arccos\dfrac{x}{\sqrt{x^2+y^2}} \end{pmatrix}, \qquad x^2+y^2>0.$$

Como la imagen de *arccos* (la rama principal) es el intervalo $0 \leq \theta \leq \pi$, esta función es la inversa de la función obtenida a partir de T restringiendo su dominio mediante la condición adicional $0 \leq \theta \leq \pi$. Para obtener valores de θ en el intervalo $\pi < \theta < 2\pi$, que corresponden a $y < 0$, sumamos π a la tercera coordenada en la fórmula dada.

La Fig. 11 muestra las tres superficies en $\mathcal{R}^3$ definidas por las ecuaciones

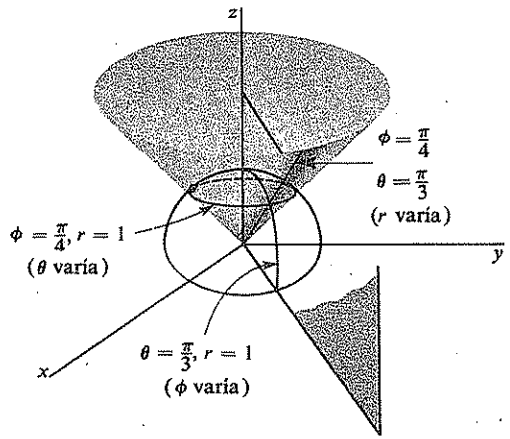


Figura 11

en coordenadas esféricas $r = 1$, $\phi = \pi/4$ y $\theta = \pi/3$, respectivamente. Las ecuaciones correspondientes en coordenadas rectangulares se obtienen de la expresión para T^{-1} antes mencionada y son

$$\begin{aligned} &x^2+y^2+z^2=1, \qquad x^2+y^2>0 \\ &z=\frac{\sqrt{2}}{2}\sqrt{x^2+y^2+z^2}, \qquad z>0 \\ &y=\sqrt{3}\,x, \qquad x>0, \end{aligned}$$

respectivamente.

Estas coordenadas se llaman «curvilíneas», ya que si todas salvo una de las coordenadas no rectangulares se mantienen fijas, al variar éstas, se obtiene una curva en $\mathcal{R}^n$. Así, en coordenadas polares, las curvas coordenadas son círculos y rectas como lo muestra la Fig. 12. En el caso de coordenadas esféricas, algunas curvas coordenadas típicas son el círculo, el semicírculo y la semirrecta obtenidas como intersecciones de las superficies indicadas en la Fig. 11. Las

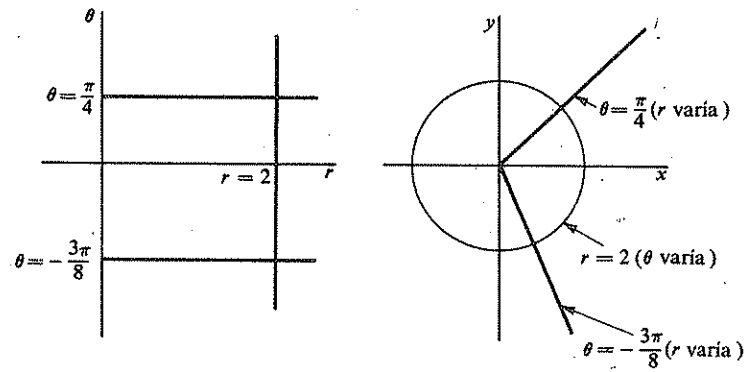


Figura 12

curvas y superficies que se obtienen haciendo variar una o más coordenadas curvilíneas juegan el mismo papel que las rectas y planos coordenados en $\mathcal{R}^n$. Por ejemplo, decir que un punto de $\mathcal{R}^3$ tiene coordenadas $(x, y, z) = (1, 2, 1)$ es lo mismo que decir que él está en la intersección de los planos $x = 1$, $y = 2$ y $z = 1$. En forma análoga, decir que un punto de $\mathcal{R}^3$ tiene coordenadas esféricas $(r, \phi, \theta) = (1, \pi/4, \pi/3)$ es lo mismo que decir que él está en la intersección de las superficies mostradas en la Fig. 11.

Generalizando a partir de los ejemplos precedentes, vemos que un sistema de coordenadas curvilíneas en $\mathcal{R}^n$ está determinado por una función $\mathcal{U}^n \xrightarrow{T} \mathcal{R}^n$. Se supone que, para algún subconjunto abierto N del dominio de T, la restricción de T a N es inyectiva y, por tanto, tiene inverso T^{-1}. Las **coordenadas**

curvilíneas, determinadas por T y N de un punto $\mathbf{x}$ en el conjunto $T(N)$ son

$$\begin{pmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{pmatrix} = T^{-1} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}.$$

Es conveniente imponer condiciones de regularidad más o menos fuertes a una transformación de coordenada. En forma específica, supondremos que, para todo $\mathbf{u}$ en N, T es continuamente diferenciable en $\mathbf{u}$ y que $T'(\mathbf{u})$ es invertible.

Los cambios de coordenadas polares y esféricas representados por

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\mathrm{sen}\,\theta \end{pmatrix}$$

y

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\,\mathrm{sen}\,\phi\cos\theta \\ r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta \\ r\cos\phi \end{pmatrix}$$

tienen las matrices jacobianas

$$\begin{pmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{pmatrix} \tag{5}$$

y

$$\begin{pmatrix} \mathrm{sen}\,\phi\cos\theta & r\cos\phi\cos\theta & -r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\phi\,\mathrm{sen}\,\theta & r\cos\phi\,\mathrm{sen}\,\theta & r\,\mathrm{sen}\,\phi\cos\theta \\ \cos\phi & -r\,\mathrm{sen}\,\phi & 0 \end{pmatrix}, \tag{6}$$

respectivamente. Las matrices (5) y (6), y más generalmente las matrices jacobianas de transformaciones coordenadas diferenciables, tienen una interpretación geométrica sencilla. Cada columna del jacobiano se obtiene derivando las funciones coordenadas con respecto de una sola variable, mientras se mantienen las otras fijas. Esto significa que la columna j de la matriz representa un vector tangente a la curva de la coordenada curvilínea que se obtiene al variar la coordenada j. Es decir, si la transformación de coordenadas está dada por $\mathfrak{U}^n \xrightarrow{T} \mathfrak{R}^n$, entonces la columna j de la matriz de la diferencial $T'(\mathbf{u}_0)$ es un vector tangente en $\mathbf{x}_0 = T(\mathbf{u}_0)$ a la curva coordenada obtenida al variar solo

la coordenada j de $\mathbf{u}_0$, denotaremos este vector por $\mathbf{c}_j$. La Fig. 13 muestra los vectores tangentes (con sus puntos iniciales trasladados al punto $\mathbf{x}_0$) de algunas

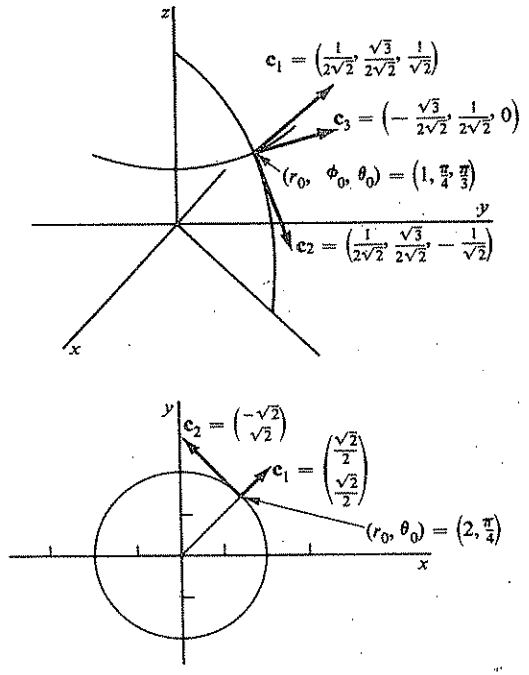


Figura 13

curvas coordenadas polares y esféricas. Como las curvas coordenadas están en el espacio cartesiano, las coordenadas de $\mathbf{c}_1, \ldots, \mathbf{c}_n$ son coordenadas rectangulares y no curvilíneas.

Demostraremos que la matriz jacobiana de una transformación de coordenadas puede ser interpretada como la matriz de una transformación lineal de coordenadas. Para ver esto, consideremos las coordenadas curvilíneas en $\mathfrak{R}^n$ dadas por $\mathbf{x} = T(\mathbf{u})$, donde $\mathbf{u}$ es la coordenada curvilínea variable. Fijemos un punto $\mathbf{x}_0$ con coordenadas curvilíneas $\mathbf{u}_0$. Podemos tomar $\mathbf{x}_0$ como nuevo origen y vectores unitarios $\mathbf{e}_1, \ldots, \mathbf{e}_n$ con la misma dirección que los vectores de la base natural de $\mathfrak{R}^n$. Entonces la multiplicación por la matriz $T'(\mathbf{u}_0)$ transforma los vectores básicos $\mathbf{e}_1, \ldots, \mathbf{e}_n$ en los vectores $\mathbf{c}_1, \ldots, \mathbf{c}_n$ tangentes a las curvas coordenadas curvilíneas. La Fig. 14 ilustra la relación entre los $\mathbf{e}_i$ y los $\mathbf{c}_i$. Nótese que los vectores $\mathbf{c}_1, \ldots, \mathbf{c}_n$ serán linealmente independientes si, y solo si, $T'(\mathbf{u}_0)$ es invertible. Esta es una razón para exigir no solo que una transformación de coordenadas sea inyectiva, sino también que, en una vecindad de un punto, su derivada sea invertible.

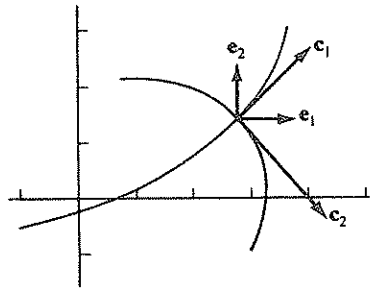


Figura 14

Ejemplo 3. (*Coordenadas cilíndricas en* $\mathfrak{R}^3$.) La transformación de coordenadas está definida por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\mathrm{sen}\,\theta \\ z \end{pmatrix}, \qquad \begin{cases} 0 < r < \infty \\ -\pi < \theta \leq \pi \\ -\infty < z < \infty. \end{cases}$$

La matriz jacobiana es

$$\begin{pmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta & 0 \\ \mathrm{sen}\,\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Las superficies coordenadas y los vectores tangentes a las curvas coordenadas se muestran en la Fig. 15. Nótese que el determinante jacobiano es

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = r.$$

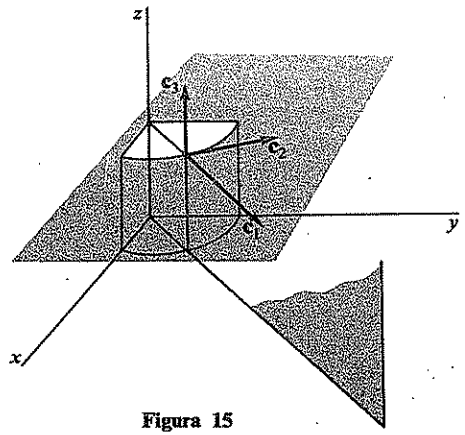


Figura 15

Los cálculos con coordenadas curvilíneas son mucho más simples si las curvas coordenadas, y los vectores $\mathbf{c}_k$, son ortogonales. (Este es el caso en los ejemplos considerados.) En ese caso, es habitual remplazar los vectores $\mathbf{c}_k$ por vectores unitarios con la misma dirección. Así, si tomamos $h_1 = |\mathbf{c}_1|, \ldots, h_n = |\mathbf{c}_n|$, tenemos la base local ortonormal $(1/h_1\mathbf{c}_1, \ldots, 1/h_n\mathbf{c}_n)$. El resultado de esto es que la matriz H, cuyas columnas son las coordenadas rectangulares de $1/h_1\mathbf{c}_1, \ldots, 1/h_n\mathbf{c}_n$, es una matriz ortogonal cuya inversa es su transpuesta.

Ejemplo 4. En el caso de coordenadas esféricas en $\mathfrak{R}^3$, tenemos $h_1 = 1$, $h_2 = r$, $h_3 = r\,\mathrm{sen}\,\phi$, de modo que

$$H = \begin{pmatrix} \mathrm{sen}\,\phi\cos\theta & \cos\phi\cos\theta & -\mathrm{sen}\,\theta \\ \mathrm{sen}\,\phi\,\mathrm{sen}\,\theta & \cos\phi\,\mathrm{sen}\,\theta & \cos\theta \\ \cos\phi & -\mathrm{sen}\,\phi & 0 \end{pmatrix}.$$

Hemos supuesto que $0 < \phi < \pi$, de modo que $\mathrm{sen}\,\phi > 0$. Entonces, como la matriz H es ortogonal, $H^{-1} = H^t$ y

$$H^{-1} = \begin{pmatrix} \mathrm{sen}\,\phi\cos\theta & \mathrm{sen}\,\phi\,\mathrm{sen}\,\theta & \cos\phi \\ \cos\phi\cos\theta & \cos\phi\,\mathrm{sen}\,\theta & -\mathrm{sen}\,\phi \\ -\mathrm{sen}\,\theta & \cos\theta & 0 \end{pmatrix}.$$

Hemos visto que los vectores $\mathbf{c}_1, \ldots, \mathbf{c}_n$, tangentes a las curvas coordenadas, describen aproximadamente la naturaleza de un sistema de coordenadas curvilíneas. Los productos interiores de estos vectores aparecen a menudo, de modo que existe una notación especial para ellos

$$\mathbf{c}_i \cdot \mathbf{c}_j = g_{ij}, \qquad i, j = 1, \ldots, n.$$

Como $\mathbf{c}_i \cdot \mathbf{c}_j = \mathbf{c}_j \cdot \mathbf{c}_i$, tenemos que $g_{ij} = g_{ji}$. Si los vectores $\mathbf{c}_1, \ldots, \mathbf{c}_n$ son ortogonales, como sucede muchas veces en la práctica, entonces solo los productos internos $\mathbf{c}_i \cdot \mathbf{c}_i = g_{ii}$ serán distintos de cero. Muchas fórmulas importantes pueden ser expresadas en coordenadas curvilíneas usando las funciones g_{ij} y sin referencia a las funciones coordenadas curvilíneas usadas.

Supongamos que están dadas coordenadas curvilíneas en el plano xy por

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix}.$$

El jacobiano de la transformación coordenada es

$$\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix},$$

de donde

$$g_{11} = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2,$$

$$g_{12} = g_{21} = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v},$$

$$g_{22} = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2.$$

Ahora, supongamos que

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$$

es una función diferenciable del intervalo $[a, b]$ en $\mathfrak{U}^2$. Entonces la ecuación

$$f(t) = \begin{pmatrix} x(u(t), v(t)) \\ y(u(t), v(t)) \end{pmatrix}, \qquad a \le t \le b,$$

es la representación paramétrica de una curva en $\mathfrak{R}^2$. El vector tangente a esta curva se calcula mediante la regla de la cadena, obteniéndose

$$\frac{df}{dt} = \begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x}{\partial u}\dfrac{du}{dt} + \dfrac{\partial x}{\partial v}\dfrac{dv}{\partial t} \\ \dfrac{\partial y}{\partial u}\dfrac{du}{dt} + \dfrac{\partial y}{\partial v}\dfrac{dv}{dt} \end{pmatrix}.$$

Entonces, el largo del vector tangente df/dt es

$$\begin{aligned} \left|\frac{df}{dt}\right| &= \left(\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right)^{1/2} \\ &= \left(\left[\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2\right]\left(\frac{du}{dt}\right)^2 + 2\left[\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v}\right]\frac{du}{dt}\frac{dv}{dt} \right. \\ &\qquad \left. + \left[\left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2\right]\left(\frac{dv}{dt}\right)^2\right)^{1/2} \\ &= \left(g_{11}\left(\frac{du}{dt}\right)^2 + 2g_{12}\frac{du}{dt}\frac{dv}{dt} + g_{22}\left(\frac{dv}{dt}\right)^2\right)^{1/2}. \end{aligned}$$

Haciendo un cálculo similar en $\mathfrak{R}^n$, se obtiene la fórmula

$$\left|\frac{df}{dt}\right| = \left(\sum_{i,j=1}^{n} g_{ij}\frac{du_i}{dt}\frac{du_j}{dt}\right)^{1/2}, \tag{7}$$

donde $f(t) = (x_1(u_1(t), \ldots, u_n(t)), \ldots, x_n(u_1(t), \ldots, u_n(t)))$.

Una vez que g_{ij} ha sido calculado, para un determinado sistema de coordenadas, la Ec. (7) puede ser usada para cualquier curva diferenciable, sustituyendo en ella las componentes del vector

$$\begin{pmatrix} \dfrac{du}{dt} \\ \dfrac{dv}{dt} \end{pmatrix}.$$

Para ser más específico, supongamos que tenemos las coordenadas polares en el plano

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\mathrm{sen}\,\theta \end{pmatrix}.$$

La matriz jacobiana es

$$\begin{pmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{pmatrix}.$$

Por tanto,

$$g_{11} = 1, \qquad g_{12} = g_{21} = 0, \qquad g_{22} = r^2.$$

Una curva $f(\theta) = \begin{pmatrix} r(\theta)\cos\theta \\ r(\theta)\,\mathrm{sen}\,\theta \end{pmatrix}$ tiene un vector tangente $df/d\theta$ de largo

$$\left|\frac{df}{d\theta}\right| = \left(\left(\frac{dr}{d\theta}\right)^2 + r^2\right)^{1/2}.$$

Ejemplo 5. Si (r, ϕ, θ) son coordenadas esféricas en $\mathfrak{R}^3$, el largo de la curva γ definida por una ecuación

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} r(t) \\ \phi(t) \\ \theta(t) \end{pmatrix}, \qquad a \le t \le b,$$

puede ser calculado usando (7). La derivada de la transformación coordenada

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\,\mathrm{sen}\,\phi\cos\theta \\ r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta \\ r\cos\phi \end{pmatrix}$$

está dada por la matriz

$$\begin{pmatrix} \operatorname{sen}\phi\cos\theta & r\cos\phi\cos\theta & -r\operatorname{sen}\phi\operatorname{sen}\theta \\ \operatorname{sen}\phi\operatorname{sen}\theta & r\cos\phi\operatorname{sen}\theta & r\operatorname{sen}\phi\cos\theta \\ \cos\phi & -r\operatorname{sen}\phi & 0 \end{pmatrix}.$$

La función g_{ij} es el producto interno de las columnas i y j de esta matriz. Por tanto,

$$\begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\operatorname{sen}^2\phi \end{pmatrix}.$$

Entonces de la Ec. (7) se tiene que

$$l(\gamma) = \int_a^b \sqrt{\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2 + r^2\operatorname{sen}^2\phi\left(\frac{d\theta}{dt}\right)^2}\, dt.$$

En particular, la curva λ definida por

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} 1 \\ t \\ t \end{pmatrix}, \qquad 0 \le t \le \frac{\pi}{2}$$

tiene largo

$$l(\lambda) = \int_0^{\pi/2} \sqrt{1 + \operatorname{sen}^2 t}\, dt.$$

Esta integral no puede ser calculada mediante una integral indefinida elemental. Sin embargo, ella puede ser aproximada por métodos sencillos. También puede ser reducida a una integral elíptica como sigue. Remplazando t por $\pi/2 - \theta$, y usando la relación $\cos^2\theta = 1 - \operatorname{sen}^2\theta$, obtenemos

$$\int_0^{\pi/2} \sqrt{1 + \operatorname{sen}^2 t}\, dt = \sqrt{2}\int_0^{\pi/2} \sqrt{1 - \tfrac{1}{2}\operatorname{sen}^2\theta}\, d\theta.$$

Esta última integral ha sido tabulada (véase, por ejemplo, R. S. Burington, *Handbook of Mathematical Tables*, Handbook Publishers, Inc., Sandusky, Ohio, 1965) y tenemos

$$l(\lambda) \approx \sqrt{2}\,(1{,}35) \approx 1{,}91.$$

EJERCICIOS

1. Dibuje las tres curvas siguientes dada en coordenadas polares por

   (a) $r = \theta$, $0 \le \theta \le \pi/2$.
   (b) $r(\operatorname{sen}\theta - \cos\theta) = \pi/2$, $\pi/2 \le \theta \le \pi$.
   (c) $r = \pi/2\cos\theta$, $\pi \le \theta \le 3\pi/2$.

2. Dibuje las curvas y superficies en $\mathcal{R}^3$ expresadas en coordenadas esféricas por

   (a) $r = 2$, $0 \le \theta \le \pi/4$, $\pi/4 \le \phi \le \pi/2$.
   (b) $1 \le r \le 2$, $\theta = \pi/2$, $\phi = \pi/4$.
   (c) $0 \le r \le 1$, $0 \le \theta \le \pi/2$, $\phi = \pi/4$.
   (d) $0 \le r \le 1$, $\theta = \pi/4$, $0 \le \phi \le \pi/4$.

3. Use coordenadas cilíndricas en $\mathcal{R}^3$ para describir la región definida en coordenadas rectangulares por $0 \le x$, $x^2 + y^2 \le 1$.

4. Sean (r, θ) coordenadas polares en $\mathcal{R}^2$. La ecuación

$$\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \operatorname{sen} t \\ t \end{pmatrix}, \qquad 0 \le t \le \frac{\pi}{2},$$

   describe una curva en $\mathcal{U}^2$. Dibuje esa curva y su imagen en $\mathcal{R}^2$ por la transformación en coordenadas polares.

5. Sean (r, ϕ, θ) las coordenadas esféricas en $\mathcal{R}^3$. La ecuación

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} 1 \\ t \\ t \end{pmatrix}$$

   determina una curva en $\mathcal{R}^3$ (y también una curva en $\mathcal{U}^3$, el espacio $r\phi\theta$). Dibuje la curva en $\mathcal{R}^3$. [*Sugestión*: La curva yace en una esfera.]

6. Demuestre que las matrices jacobianas (5) y (6) de las transformaciones de coordenadas polar y esférica, dadas en los Ejemplos 2 y 3, tienen inversos.

7. Las ecuaciones

$$\left.\begin{aligned} x &= ar\operatorname{sen}\phi\cos\theta \\ y &= br\operatorname{sen}\phi\operatorname{sen}\theta \\ z &= cr\cos\phi \end{aligned}\right\}, \qquad a, b, c > 0,$$

   definen coordenadas elipsoidales en $\mathcal{R}^3$. Dibuje un ejemplo típico de cada una de las tres superficies coordenadas si $a = 1$, $b = c = 2$.

8. Calcule las componentes cartesianas de los vectores tangentes a las curvas coordenadas para las coordenadas elipsoidales dadas en el Ejercicio 7, cuando $a = b = 1$, $c = 0$ y $r = \frac{1}{2}$, $\phi = \theta = \pi/2$.

9. Sean r, ϕ, θ coordenadas esféricas en $\mathcal{R}^3$. La ecuación

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix}$$

determina una curva en $\mathcal{R}^3$. Calcule las componentes cartesianas del vector tangente a la curva.

10. Demuestre que en las coordenadas esféricas tridimensionales del Ejemplo 2, la esfera $x_1^2 + x_2^2 + x_3^2 = 1$ tiene la ecuación $r = 1$.

11. Las coordenadas elípticas en el plano están determinadas por

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ar\cos\theta \\ br\,\text{sen}\,\theta \end{pmatrix}, \qquad a > 0, \quad b > 0.$$

(a) Calcule los coeficientes g_{ij} para este sistema de coordenadas.
(b) ¿Para cuáles a y b será siempre verdad que $g_{ij} = 0$ si $i \neq j$?

12. Verifique la afirmación hecha más arriba que la columna j de la matriz jacobiana en $\mathbf{u}_0$ de la transformación coordenada $\mathcal{U}^n \xrightarrow{T} \mathcal{R}^n$ es un vector tangente en $\mathbf{x}_0 = T(\mathbf{u}_0)$ a la curva coordenada obtenida haciendo variar la coordenada j de $\mathbf{u}_0$.

13. Si $f(x, y) = f(r, \theta)$, donde $(x, y) = (r\cos\theta,\ r\,\text{sen}\,\theta)$, demuestre que

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r}\frac{\partial f}{\partial r}.$$

14. (a) Encuentre la fórmula para el largo de arco de una curva en el plano determinada en coordenadas polares por

$$\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r(t) \\ \theta(t) \end{pmatrix}, \qquad a \leq t \leq b.$$

(b) Calcule el largo de la curva $\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 2t \\ t \end{pmatrix}$, $\quad 0 \leq t \leq 2.$

[*Resp*: $2\sqrt{5} + \log(2 + \sqrt{5})$.]

(c) Dibuje la curva.

15. Una ecuación

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}, \qquad a \leq t \leq b,$$

determina una curva γ en la superficie canónica

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u\cos v\,\text{sen}\,\alpha \\ u\,\text{sen}\,v\,\text{sen}\,\alpha \\ u\cos\alpha \end{pmatrix}, \qquad \begin{cases} 0 \leq u \leq \infty, \\ 0 \leq v \leq 2\pi, \end{cases}$$

en $\mathcal{R}^3$, donde α es fijo, $0 < \alpha < \pi/2$. Encuentre la fórmula general para el largo de arco de γ.

16. Si $T(u_1, \ldots, u_n) = (x_1, \ldots, x_n)$ define un sistema de coordenadas curvilíneas en una región D de $\mathcal{R}^n$

(a) Demuestre que $g_{ij} = \sum_{k=1}^{n} \frac{\partial x_k}{\partial u_i}\frac{\partial x_k}{\partial u_j}$, para $i, j = 1, \ldots, n$.

(b) Si (g^{ij}) es la matriz inversa de (g_{ij}), demuestre que

$$g^{ij} = \sum_{k=1}^{n} \frac{\partial u_i}{\partial x_k}\frac{\partial u_j}{\partial x_k}, \qquad i, j = 1, \ldots, n.$$

(c) Demuestre que si $f(u_1, \ldots, u_n)$ representa una función diferenciable en una región D en $\mathcal{R}^n$, entonces

$$\nabla f = \sum_{j=1}^{n} \left[\sum_{i=1}^{n} \frac{\partial f}{\partial u_i} g^{ij}\right] \mathbf{c}_j.$$

(d) Demuestre que si T define un sistema de coordenadas ortogonales en D, entonces

$$\nabla f = \sum_{i=1}^{n} \frac{1}{h^2}\frac{\partial f}{\partial u_i}\mathbf{c}_i.$$

6. Teoremas de la función inversa y de la función implícita

Si se considera que una función f transforma vectores $\mathbf{x}$ en vectores $\mathbf{y}$ de la imagen de f, es natural partir de $\mathbf{y}$ y preguntarse cuáles vectores $\mathbf{x}$ son transformados por f en $\mathbf{y}$. En particular, podemos preguntarnos si existe una función que invierte la acción de f. Si existe una función f^{-1} con la propiedad de que $f^{-1}(\mathbf{y}) = \mathbf{x}$ si, y solo si, $f(\mathbf{x}) = \mathbf{y}$, entonces f^{-1} se llama la **función inversa** de f. El dominio de f^{-1} es la imagen de f y la imagen de f^{-1} es el dominio de f. Algunos ejemplos familiares de funciones y sus inversas son

$$\begin{cases} f(x) = x^2, & x \geq 0. \\ f^{-1}(y) = \sqrt{y}, & y \geq 0. \end{cases}$$

$$\begin{cases} f(x) = e^x, & -\infty < x < \infty, \\ f^{-1}(y) = \ln y, & y > 0. \end{cases}$$

$$\begin{cases} f(x) = \text{sen}\,x, & -\dfrac{\pi}{2} \leq x \leq \dfrac{\pi}{2}. \\ f^{-1}(y) = \text{arcsen}\,y, & -1 \leq y \leq 1. \end{cases}$$

La función inversa f^{-1} no debe ser confundida con la recíproca $1/f$. Por ejemplo, si $f(\mathbf{x}) = x^2$, entonces $f^{-1}(2) = \sqrt{2}$, mientras que $(f(2))^{-1} = 1/f(2) = \frac{1}{4}$.

Recordemos que una función es inyectiva si cada elemento de su imagen proviene de exactamente un elemento del dominio. Un hecho que se usa repetidamente es que una función tiene inverso si, y solo si, ella es inyectiva.

La función inversa L^{-1} de una función lineal invertible $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^m$ es lineal. En efecto, aplicando L a ambos miembros, vemos que

$$L^{-1}(a\mathbf{y}_1 + b\mathbf{y}_2) = aL^{-1}(\mathbf{y}_1) + bL^{-1}(\mathbf{y}_2),$$

cuando $\mathbf{y}_1$ e $\mathbf{y}_2$ están en la imagen de L. Si la dimensión de $\mathcal{R}^n$ es menor que la de $\mathcal{R}^m$, la imagen de L es un subespacio propio de $\mathcal{R}^m$. En este caso L^{-1} no está definida en todo $\mathcal{R}^m$. Por otra parte, si $\mathcal{R}^n$ y $\mathcal{R}^m$ tienen la misma dimensión, el dominio de L^{-1} es todo $\mathcal{R}^m$. Por tanto, la función inversa de una función lineal inyectiva $\mathcal{R} \xrightarrow{F} \mathcal{R}^n$ es una función lineal $\mathcal{R}^n \xrightarrow{L^{-1}} \mathcal{R}^n$.

Ejemplo 1. Consideremos la función afín $\mathcal{R}^3 \xrightarrow{A} \mathcal{R}^3$ definida por

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix}\begin{pmatrix} x-1 \\ y-0 \\ z-1 \end{pmatrix} + \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix}.$$

Es obvio que una función afín $A(\mathbf{x}) = L(\mathbf{x} - \mathbf{x}_0) + \mathbf{y}_0$ es inyectiva si, y solo si, la función lineal L lo es. En este ejemplo,

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{y} \quad L(\mathbf{x}) = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

La matriz inversa de

$$\begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix}$$

se puede calcular y es

$$\begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix}.$$

De esto sigue que L y, por tanto, A, tienen inversos. En efecto, si $A(\mathbf{x}) = \mathbf{y}$, entonces

$$A(\mathbf{x}) = L(\mathbf{x} - \mathbf{x}_0) + \mathbf{y}_0$$

y

$$A^{-1}(\mathbf{y}) = L^{-1}(\mathbf{y} - \mathbf{y}_0) + \mathbf{x}_0. \tag{1}$$

Sustituyendo $\mathbf{y}$ por $A(\mathbf{x})$, comprobamos que ésta es la expresión correcta para A^{-1}. Tenemos

$$A^{-1}(A(\mathbf{x})) = L^{-1}(L(\mathbf{x} - \mathbf{x}_0)) + \mathbf{x}_0 = \mathbf{x}.$$

Por tanto,

$$A^{-1}\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix}\begin{pmatrix} u-1 \\ v-5 \\ w-2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Es evidente que este método nos permite encontrar la inversa de cualquier transformación afín $\mathcal{R}^n \xrightarrow{A} \mathcal{R}^n$, siempre que ésta exista.

Tenemos los siguientes criterios para decidir si una función lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^m$ tiene un inverso. Si M es la matriz de L, entonces por el Teorema 4.2, las columnas de M son los vectores $L(\mathbf{e}_j)$ y, por tanto, generan la imagen de L. Entonces, L es inyectiva y tiene inversa si, y solo si, las columnas de M son linealmente independientes. Por otra parte, si *M es una matriz cuadrada*, L tiene un inverso si, y solo si, existe la matriz inversa M^{-1}. Recordamos que M^{-1} existe si, y solo si, $\det M \neq 0$.

El propósito principal de esta sección es el estudio de las inversas de funciones vectoriales no lineales. Dada una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$, uno puede preguntar: (1) ¿Tiene f una inversa? y (2) Si tiene una, ¿cuáles son sus propiedades? En general, no es fácil contestar estas preguntas por simple consideración de la función. Por otra parte, sabemos determinar si una función afín tiene o no inversa, y cómo calcularla si ella existe. Además, si f es diferenciable en un punto $\mathbf{x}_0$, cerca de ese punto f puede ser aproximada por una función afín A. Por esa razón, se puede conjeturar que si se restringe el dominio de f a puntos cercanos a $\mathbf{x}_0$, entonces f tendrá una inversa si A la tiene. Por último, se puede esperar que A^{-1} es la función afín que aproxima a f^{-1} cerca de $f(\mathbf{x}_0)$. Salvo detalles, estas afirmaciones son verdaderas y forman el contenido del teorema de la función inversa.

6.1 Teorema de la función inversa

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ una función continuamente diferenciable tal que $f'(\mathbf{x}_0)$ tiene inverso. Entonces existe un conjunto abierto N que contiene a $\mathbf{x}_0$ tal que cuando se restringe f a N, ella tiene una inversa continuamente diferenciable f^{-1}. El conjunto $f(N)$ es abierto. Además,

$$[f^{-1}]'(\mathbf{y}_0) = [f'(\mathbf{x}_0)]^{-1},$$

donde $\mathbf{y}_0 = f(\mathbf{x}_0)$. Es decir, la diferencial de la función en $\mathbf{y}_0$ es la inversa de la diferencial de f en $\mathbf{x}_0$.

La existencia de f^{-1} se demostrará en el Apéndice. Una vez que se ha es-

tablecido la existencia, podemos escribir f^{-1} o $f = I$, donde $\mathcal{R}^n \xrightarrow{I} \mathcal{R}^n$ es la transformación identidad en la vecindad N. Entonces, por la regla de la cadena, tenemos que

$$[f^{-1}]'(\mathbf{y}_0)f'(\mathbf{x}_0) = I \quad \text{o} \quad [f^{-1}]'(\mathbf{y}_0) = [f'(\mathbf{x}_0)]^{-1}.$$

puesto que la transformación identidad es su propia diferencial.

Para funciones reales de una variable, no es difícil demostrar la existencia de una función inversa. Supongamos que $\mathcal{R} \xrightarrow{f} \mathcal{R}$ satisface la condición de diferenciabilidad del teorema y supongamos que $f'(\mathbf{x}_0) \neq 0$, el significado geométrico de este hecho es que el gráfico de f no debe tener una tangente horizontal. Para ser más específicos, supongamos que $f'(x_0) > 0$. Como f' es

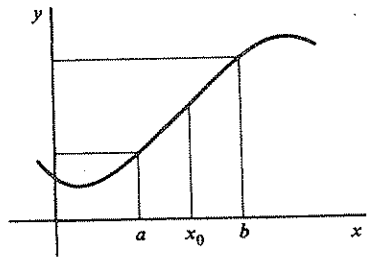


Figura 16

continua, tenemos que $f'(x) > 0$ para todo x en algún intervalo $a < x < b$ que contiene a x_0, como el indicado en la Fig. 16. Afirmamos que f es inyectiva en este intervalo. En efecto, sean x_1 y x_2 dos puntos en el intervalo tales que $x_1 < x_2$. Por el teorema del valor medio, tenemos que

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_0),$$

para algún x_0 en el intervalo $x_1 < x < x_2$. Como $f'(x_0) > 0$ y $x_2 - x_1 > 0$, tenemos que

$$f(x_2) - f(x_1) > 0.$$

Es decir, f es estrictamente creciente en el intervalo $a < x < b$ y nuestra afirmación queda demostrada. Por tanto, f restringida a este intervalo tiene una inversa. Las otras conclusiones del teorema de la función inversa también pueden ser demostradas en forma sencilla en este caso particular.

Ejemplo 2. Consideremos la función f definida por

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^3 - 2xy^2 \\ x + y \end{pmatrix}, \quad \begin{matrix} -\infty < x < \infty, \\ -\infty < y < \infty. \end{matrix}$$

En el punto

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

la diferencial $d_{\mathbf{x}_0}f$ está definida por la matriz jacobiana

$$\begin{pmatrix} 3x^2 - 2y^2 & -4xy \\ 1 & 1 \end{pmatrix}_{x=1,\,y=-1} = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}.$$

La inversa de esta matriz es

$$\begin{pmatrix} -\frac{1}{8} & \frac{4}{8} \\ \frac{1}{8} & -\frac{1}{8} \end{pmatrix}.$$

Como f es continuamente diferenciable, del teorema de la función inversa concluimos que en algún conjunto abierto que contiene a $\mathbf{x}_0$, f tiene una inversa f^{-1}. Además, si

$$\mathbf{y}_0 = f(\mathbf{x}_0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

entonces la matriz de la diferencial $d_{\mathbf{y}_0}f^{-1}$ es

$$\begin{pmatrix} -\frac{1}{8} & \frac{4}{8} \\ \frac{1}{8} & -\frac{1}{8} \end{pmatrix}.$$

A pesar de que sería difícil calcular f^{-1} explícitamente, es fácil encontrar la transformación afín que aproxima a f^{-1} en una vecindad del punto $\mathbf{y}_0$. Ella es la inversa A^{-1} de la función afín A que aproxima a f cerca de $\mathbf{x}_0$. Usando el teorema de la función o la Ec. (1) del Ejemplo 1, tenemos que

$$\begin{aligned} A(\mathbf{x}) &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= \mathbf{y}_0 + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0). \\ A^{-1}(\mathbf{y}) &= f^{-1}(\mathbf{y}_0) + [f^{-1}]'(\mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) \\ &= \mathbf{x}_0 + [f'(\mathbf{x}_0)]^{-1}(\mathbf{y} - \mathbf{y}_0). \end{aligned}$$

Por tanto, si hacemos $\mathbf{y} = \begin{pmatrix} u \\ v \end{pmatrix}$, tenemos

$$\begin{aligned} A^{-1}\begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -\frac{1}{8} & \frac{4}{8} \\ \frac{1}{8} & -\frac{1}{8} \end{pmatrix}\begin{pmatrix} u+1 \\ v-0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{8} & \frac{4}{8} \\ \frac{1}{8} & -\frac{1}{8} \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \frac{2}{8} \\ -\frac{2}{8} \end{pmatrix}. \end{aligned}$$

Ejemplo 3. Las ecuaciones

$$u = x^4y + x, \qquad v = x + y^3$$

definen una transformación de $\mathcal{R}^2$ en $\mathcal{R}^2$. La matriz de la diferencial de esta transformación en $(x, y) = (1, 1)$ es

$$\begin{pmatrix} 4x^3y + 1 & x^4 \\ 1 & 3y^2 \end{pmatrix}_{(x,y)=(1,1)} = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}.$$

Como las columnas de esta matriz son independientes, la diferencial tiene una inversa y, de acuerdo con el teorema de la función inversa, la transformación tiene una función inversa en una vecindad de $(x, y) = (1, 1)$. La transformación inversa estaría dada por ecuaciones de la forma

$$x = F(u, v), \qquad y = G(u, v).$$

El cálculo de F y G es difícil, pero podemos calcular fácilmente las derivadas parciales de F y G con respecto a u y v en el punto $(u, v) = (2, 2)$ que corresponde a $(x, y) = (1, 1)$. Estas derivadas parciales aparecen en las matrices jacobianas de F y G o, lo que es lo mismo, en la matriz inversa de la diferencial de la función dada. Tenemos

$$\begin{pmatrix} \dfrac{\partial F}{\partial u}(2,2) & \dfrac{\partial F}{\partial v}(2,2) \\ \dfrac{\partial G}{\partial u}(2,2) & \dfrac{\partial G}{\partial v}(2,2) \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{3}{14} & -\dfrac{1}{14} \\ -\dfrac{1}{14} & \dfrac{5}{14} \end{pmatrix}.$$

Supongamos que $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ es una función tal que las hipótesis del teorema de la función inversa son válidas en un punto $\mathbf{x}_0$. Es importante darse cuenta que el teorema no resuelve si existe o no una inversa f, solo dice que ella existe cuando se restringe f a una vecindad de $\mathbf{x}_0$. Por ejemplo, la transformación

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \end{pmatrix}, \qquad 0 < u,$$

tiene la matriz jacobiana

$$\begin{pmatrix} \cos v & -u \operatorname{sen} v \\ \operatorname{sen} v & u \cos v \end{pmatrix}$$

con matriz inversa

$$\begin{pmatrix} \cos v & \operatorname{sen} v \\ -\dfrac{1}{u}\operatorname{sen} v & \dfrac{1}{u}\cos v \end{pmatrix}.$$

Esta matriz inversa existe para todo (u, v) tal que $u > 0$. Sin embargo, si no restringimos esta transformación, la inversa no existe, ya que cuando v se aumenta en 2π, se obtiene la misma imagen. La Fig. 17 muestra dos regiones que se corresponden mediante la transformación. Si la transformación se restringe

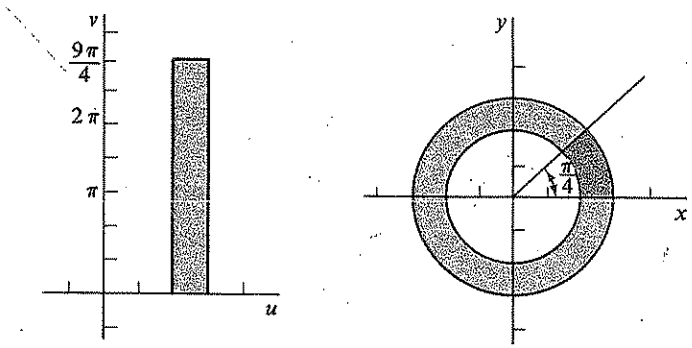


Figura 17

de modo que, por ejemplo, $0 < v < 2\pi$, entonces ella es inyectiva y tiene inverso.

En la Sec. 4 consideramos el problema de encontrar las derivadas de una función f definida implícitamente bajo la hipótesis que f satisface una ecuación $F(\mathbf{x}, f(\mathbf{x})) = 0$, con f y F diferenciables. Vimos que, para encontrar $f'(\mathbf{x}_0)$ por métodos matriciales, era necesario que $F_y(\mathbf{x}_0, f(\mathbf{x}_0))$ tuviera un inverso. Es natural que esta misma condición aparezca en el teorema siguiente sobre la existencia de una función diferenciable f definida implícitamente por F. Una demostración del teorema depende del teorema de la función inversa y la daremos en la Sec. 5 del Apéndice.

6.2 Teorema de la función implícita

Sea $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ una función continuamente diferenciable. Supongamos que para algún $\mathbf{x}_0$ en $\mathcal{R}^n$ e $\mathbf{y}_0$ en $\mathcal{R}^m$,

1. $F(\mathbf{x}_0, \mathbf{y}_0) = 0$.
2. $F_y(\mathbf{x}_0, \mathbf{y}_0)$ tiene una inversa.

Entonces existe una función continuamente diferenciable $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ definida en una vecindad M de $\mathbf{x}_0$ tal que $f(\mathbf{x}_0) = \mathbf{y}_0$ y $F(\mathbf{x}, f(\mathbf{x})) = 0$, para todo $\mathbf{x}$ en N.

Como lo viéramos en el Teorema 4.2, la derivada de f está dada por

$$f'(\mathbf{x}) = -F_y^{-1}(\mathbf{x}, f(\mathbf{x}))F_x(\mathbf{x}, f\mathbf{x}).$$

Ejemplo 4 La ecuación $x^3y + y^3x - 2 = 0$ define $y = f(x)$ implícitamente en una vecindad de $x = 1$ si $f(1) = 1$. Como función de y, $x^3y + y^3x - 2$ tiene jacobiano $(1 + 3y^2)$ en $x = 1$, el cual es invertible en $y = 1$, es decir,

$$1 + 3y^2|_{y=1} = 4 \neq 0.$$

La solución puede ser calculada por métodos habituales y es

$$y = \sqrt[3]{\frac{1}{x} + \frac{1}{x}\sqrt{\frac{x^{10} + 27}{27}}} + \sqrt[3]{\frac{1}{x} - \frac{1}{x}\sqrt{\frac{x^{10} + 27}{27}}}.$$

(Véase R. S. Burington, *Mathematical Tables and Formulas*, Handbook Publishers, Inc., Sandusky, Ohio, 1965.)

Ejemplo 5. Las ecuaciones

$$z^3x + w^2y^3 + 2xy = 0, \qquad xyzw - 1 = 0 \tag{2}$$

pueden ser escritas en la forma $F(\mathbf{x}, \mathbf{y}) = 0$, donde $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} z \\ w \end{pmatrix}$, y

$$F(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} z^3x + w^2y^3 + 2xy \\ xyzw - 1 \end{pmatrix}.$$

Sean $\mathbf{x}_0 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ y $\mathbf{y}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Entonces

$$F(\mathbf{x}_0, \mathbf{y}) = \begin{pmatrix} -z^3 - w^2 + 2 \\ zw - 1 \end{pmatrix}$$

y la matriz $F_{\mathbf{y}}(1, 1)$ es

$$\begin{pmatrix} -3z^2 & 2w \\ w & z \end{pmatrix}_{\binom{z}{w}=\binom{1}{1}} = \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}.$$

La matriz inversa existe y es

$$\begin{pmatrix} -1 & -2 \\ 2 & 3 \end{pmatrix}.$$

Como consecuencia del teorema de la función implícita, las Ecs. (2) definen implícitamente una función f en un conjunto abierto que contiene a $\mathbf{x}_0$ y tal que $f(\mathbf{x}_0) = \mathbf{y}_0$. Es decir, tenemos que

$$\begin{pmatrix} z \\ w \end{pmatrix} = f\begin{pmatrix} x \\ y \end{pmatrix},$$

de modo que z y w son funciones de x e y cerca de

$$\begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

EJERCICIOS

1. ¿Es posible que una función tenga dos inversas?
2. Demuestre que $(f^{-1})^{-1} = f$.
3. ¿Cuáles de las funciones que siguen tienen inversa?
   (a) $y = \cosh x$, $-\infty < x < \infty$.
   (b) $y = \cosh x$, $0 \leq x < \infty$.
   (c) $f(x) = \tan x$, $\begin{cases} \frac{7}{4}\pi \leq x < \frac{11}{4}\pi, \\ x \neq \frac{10}{4}\pi. \end{cases}$
   (d) $f(x) = \tan x$, $0 \leq x \leq \pi/4$.
   (e) $y = x^2 - 2x + 1$, $1 \leq x < \infty$.
   (f) $y = x^2 - 3x + 2$, $0 \leq x < \infty$.
4. Calcule A^{-1} para las siguientes funciones afines:
   (a) $A(x) = 7x + 2$.
   (b) $A\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} u - 1 \\ v - 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

   $$\left[\textit{Resp: } A^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} x - 3 \\ y - 4 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \right]$$

5. Usando el teorema de la función inversa, demuestre que las siguientes funciones tienen inversas cuando se las restringe a algún conjunto abierto que contiene a $\mathbf{x}_0$.
   (a) $f(x) = \tan x$, $x_0 = \pi/6$.
   (b) $y = x^2 - 3x + 2$, $x_0 = 4$.
   (c) $y = x^3 - 7x + 6$, $x_0 = 4$.
   (d) $f(x) = \int_{-\infty}^{x} e^{-t^2}\,dt$, $x_0 = 0$.
6. Sea $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}$.
   (a) Demuestre que para todo punto $\mathbf{x}_0$ excepto

   $$\mathbf{x}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

   la restricción de f a algún conjunto abierto que contiene a $\mathbf{x}_0$ tiene una inversa.
   (b) Demuestre que si no se restringe su dominio, f no tiene inversa.
   (c) Si f^{-1} es la inversa de f en una vecindad del punto $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, calcule la transforma-

ción afín que aproxima a f^{-1} cerca de

$$f\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}-3\\4\end{pmatrix}.$$

$$\left[\textit{Resp}: \begin{pmatrix} \frac{1}{10} & \frac{1}{5} \\ -\frac{1}{5} & \frac{1}{10} \end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix} + \begin{pmatrix}\frac{1}{2}\\1\end{pmatrix}.\right]$$

7. Encuentre la función afín que aproxima mejor a la inversa de la función

$$f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x^3 + 2xy + y^2\\ x^2 + y\end{pmatrix}$$

cerca del punto $f\begin{pmatrix}1\\1\end{pmatrix}$. Nótese que sería difícil encontrar la inversa

$$\left[\textit{Resp}: \begin{pmatrix} -\frac{1}{3} & \frac{4}{3} \\ \frac{2}{3} & -\frac{5}{3} \end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix} + \begin{pmatrix}-\frac{1}{3}\\ \frac{5}{3}\end{pmatrix}.\right]$$

8. (a) Sea T definida por

$$\begin{pmatrix}x\\y\end{pmatrix} = T\begin{pmatrix}r\\\theta\end{pmatrix} = \begin{pmatrix}r\cos\theta\\ r\,\text{sen}\end{pmatrix}, \qquad \begin{cases}0 < r,\\ 0 \le \theta < 2\pi.\end{cases}$$

Encuentre $d_{\mathbf{u}} T$ y su inversa para los puntos

$$\mathbf{u} = \begin{pmatrix}r\\\theta\end{pmatrix}$$

para los cuales ella existe.

(b) Sea S definida por

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = S\begin{pmatrix}r\\\phi\\\theta\end{pmatrix} = \begin{pmatrix}r\,\text{sen}\,\phi\cos\theta\\ r\,\text{sen}\,\phi\,\text{sen}\,\theta\\ r\cos\phi\end{pmatrix}, \qquad \begin{cases}0 < r,\\ 0 < \phi < \pi/2,\\ 0 < \theta < 2\pi.\end{cases}$$

Encuentre $d_{\mathbf{u}} S$ y su inversa para los puntos

$$\mathbf{u} = \begin{pmatrix}r\\\phi\\\theta\end{pmatrix}$$

para los cuales ella existe.

(c) Encuentre una representación explícita de S^{-1}.

9. Suponga que la función T definida por

$$\begin{pmatrix}u\\v\end{pmatrix} = T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}f(x,y)\\ g(x,y)\end{pmatrix}$$

tiene una función inversa diferenciable S definida por

$$\begin{pmatrix}x\\y\end{pmatrix} = S\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}h(u,v)\\ k(u,v)\end{pmatrix}.$$

Si $f(1,2) = 3$, $g(1,2) = 4$ y $T'(1,2)$ es igual a

$$\begin{pmatrix}3 & 5\\ 4 & 7\end{pmatrix}, \quad \text{encuentre } \frac{\partial h}{\partial v}(3,4). \qquad [\textit{Resp.} \quad -5.]$$

10. Si

$$\begin{cases}x = u + v + w,\\ y = u^2 + v^2 + w^2,\\ z = u^3 + v^3 + w^3,\end{cases}$$

calcule $\partial v/\partial y$ en la imagen $(x, y, z) = (2, 6, 8)$ de $(u, v, w) = (1, 2, -1)$.

[*Resp.* 0.]

11. Sea

$$f\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}u^2 + u^2v + 10v\\ u + v^3\end{pmatrix}.$$

(a) Demuestre que f tiene una inversa f^{-1} en una vecindad del punto $\begin{pmatrix}1\\1\end{pmatrix}$.

(b) Encuentre un valor aproximado de $f^{-1}\begin{pmatrix}11{,}8\\ 2{,}2\end{pmatrix}$.

12. ¿Tiene una inversa la función $f(t) = \begin{pmatrix}t\\t\\t\end{pmatrix}$?

13. Demuestre que la función diferenciable

$$F(x,y,z) = \begin{pmatrix} f(x,y,z)\\ g(x,y,z)\\ f(x,y,z) + g(x,y,z)\end{pmatrix}$$

no puede tener una inversa diferenciable.

14. A pesar de que la existencia de una inversa de $d_{x_0} f$ es necesaria para demostrar el teorema de la función inversa, es enteramente posible que esta condición falle a pesar de que exista la inversa. Verifique este hecho usando el ejemplo $f(x) = x^3$.

15. El teorema de la función inversa es la modificación correcta de la afición simple, pero falsa, que si $d_x f$ tiene una inversa, entonces f también la tiene. La afirmación conversa, que dice que si f tiene una inversa, $d_x f$ también tiene inversa, es también falsa

(véase Ejercicio 14). Sin embargo, ella también es casi verdadera. Usando la regla de la cadena, demuestre la forma corregida: *Si f es diferenciable y tiene una inversa diferenciable, entonces $d_x f$ es inyectiva.*

16. Considere la función $\mathbb{R} \xrightarrow{f} \mathbb{R}$ definida por

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \operatorname{sen} \frac{1}{x}, & \text{si } x \neq 0. \\ 0, & \text{si } x = 0. \end{cases}$$

Demuestre que $d_0 f$ es inyectiva, pero que f no tiene inversa en una vecindad de $x = 0$. ¿Qué dice este ejemplo acerca del teorema de la función inversa?

17. ¿Cuál es la función inversa de la función lineal

$$L\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ u - v \\ u \end{pmatrix}?$$

18. El teorema de la función inversa puede ser generalizado como sigue:

Sea $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$, donde $n < m$, una función continuamente diferenciable. Si $d_{x_0} f$ es inyectiva, entonces existe un conjunto abierto N que contiene a $\mathbf{x}_0$ tal que f, restringida a N, tiene una inversa f^{-1}.

Una diferencia importante entre este teorema y el de la función inversa es que aquí la imagen $f(N)$ no es un subconjunto abierto de $\mathbb{R}^m$. Una consecuencia de esto es que f^{-1} no es diferenciable en $f(\mathbf{x}_0)$.

(a) Si

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ (u + v)^2 \\ (u + v)^3 \end{pmatrix},$$

¿para cuáles puntos (u_0, v_0) tiene f una inversa en una vecindad de $f(u_0, v_0)$?

(b) Demuestre el teorema generalizado de la función inversa. [*Indicación*: Sea $\mathbf{y}_1, \ldots, \mathbf{y}_n$ una base para la imagen de $d_{x_0} f$. Extiéndala a una base $\mathbf{y}_1, \ldots, \mathbf{y}_n, \mathbf{y}_{n+1}, \ldots, \mathbf{y}_m$ de $\mathbb{R}^m$ y defina $\mathbb{R}^m \xrightarrow{g} \mathbb{R}^n$ por

$$g\left(\sum_{i=1}^{m} a_i \mathbf{y}_i\right) = (a_1, \ldots, a_n).$$

Demuestre que $(g \circ f)$ y $d_{x_0}(g \circ f) = (d_{f(x_0)} g) \circ (d_{x_0} f)$ satisfacen las condiciones del teorema de la función inversa.]

19. Considere la ecuación $(x - 2)^3 y + x e^{y-1} = 0$.

(a) ¿Está y definido implícitamente como una función de x en una vecindad de $(x, y) = (1, 1)$?

(b) ¿En una vecindad de $(0, 0)$?

(c) ¿En una vecindad de $(2, 1)$?

20. El punto $(x, y, t) = (0, 1, -1)$ satisface las ecuaciones

$$xyt + \operatorname{sen} xyt = 0, \qquad x + y + t = 0.$$

¿Están x e y definidas implícitamente como funciones de t en una vecindad de $(0, 1, -1)$?

21. La condición 2 en el teorema de la función implícita que dice que $F_y(\mathbf{x}_0, \mathbf{y}_0)$ debe tener un inverso, no es una condición necesaria para que la ecuación $F(\mathbf{x}, \mathbf{y}) = 0$ defina una función diferencial única f tal que $f(\mathbf{x}_0) = \mathbf{y}_0$. Demuestre esto tomando $F(x, y) = x^9 - y^3$ y $(x_0, y_0) = (0, 0)$.

22. Demuestre que si N' es un subconjunto abierto de $\mathbb{R}^{n+m}$, que contiene al punto $(\mathbf{x}_0, \mathbf{y}_0)$, entonces el subconjunto N de todos los $\mathbf{x}$ en $\mathbb{R}^n$ tales que $(\mathbf{x}, \mathbf{y}_0)$ están en N' es un subconjunto abierto de $\mathbb{R}^n$.

23. Demuestre que bajo las hipótesis del Teorema 6.2, existe una sola función f definida por $F(\mathbf{x}, f(\mathbf{x})) = 0$ en una vecindad de $\mathbf{x}_0$ y tal que $f(\mathbf{x}_0) = \mathbf{y}_0$. [*Indicación*: Use la función $H(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$.]

7. Superficies y tangentes

A pesar de que las representaciones explícitas, implícitas y paramétricas de superficies han sido usadas hasta ahora a modo de ilustración, todavía no hemos dado una definición precisa del término «superficie». En está sección definiremos una superficie lisa en términos de cada una de las tres representaciones y daremos una definición unificada de la noción de tangente en cada una de estas formas de presentación, mostrando cómo éstas están relacionadas. En particular, veremos para cada una de las maneras de representar una superficie —como imagen, como gráfico o como conjunto de nivel— se obtiene una representación de una tangente tomando la imagen, gráfico o conjunto de nivel de la aproximación afín de la función dada.

Un **plano** n-dimensional en $\mathbb{R}^n$ es un subespacio vectorial de dimensión n (es decir, el conjunto generado por n vectores linealmente independientes) o un subespacio afín (es decir, el trasladado de un subespacio vectorial por un vector fijo $\mathbf{y}_0$). Cuando $n = 1$ o $n = 2$, obtenemos una recta o un plano ordinario, respectivamente. Una representación paramétrica de un plano n-dimensional en $\mathbb{R}^m$ se obtiene considerando la imagen de una transformación afín $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$, donde $A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0$, con L una transformación lineal definida en $\mathbb{R}^n$. Para asegurarnos que la imagen de A sea n-dimensional, podemos exigir que la matriz de L tenga n columnas linealmente independientes, puesto que las columnas $L(\mathbf{e}_1)$ generan la imagen de L.

Superficies definidas paramétricamente

Sea S un conjunto definido paramétricamente por una función $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$. De acuerdo con la definición de la Sec. 1 del Cap. 3, esto significa que S es la imagen por f del dominio de f. Luego, restringimos f a una vecindad de $\mathbf{x}_0$ en la cual f es inyectiva. Si f es diferenciable en $\mathbf{x}_0$, entonces la función afín A que aproxima a f cerca de $\mathbf{x}_0$ está dada por $A(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ para

todo $\mathbf{x}$ en $\mathcal{R}^n$. Entonces, si A define paramétricamente un plano n-dimensional, este plano se llama el **plano tangente** a S en $f(\mathbf{x}_0)$. Nótese que se requiere que la dimensión n del plano sea igual a la dimensión del espacio dominio de f. Si, además, f es continuamente diferenciable en su dominio, el conjunto S se llama una **superficie lisa** (o **curva lisa** si $n = 1$) en todo punto en el cual existe una tangente.

Ejemplo 1. Consideremos la superficie S en el espacio tridimensional $\mathcal{R}^3$ definido paramétricamente por

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u\cos v \\ u\,\mathrm{sen}\, v \\ v \end{pmatrix}, \qquad \begin{cases} 0 \le u \le 4, \\ 0 \le v \le 2\pi. \end{cases}$$

Esta función ha sido discutida en el Ejemplo 7 de la Sec. 1 del Cap. 3 y su imagen es la superficie S mostrada en la Fig. 6. En $(u_0, v_0) = (2, \pi/2)$, la matriz de la diferencial es

$$\begin{pmatrix} \cos v_0 & -u_0 \,\mathrm{sen}\, v_0 \\ \mathrm{sen}\, v_0 & u_0 \cos v_0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

La función afín $A(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ que aproxima a f cerca de $\mathbf{x}_0$ está dada por

$$A\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ \frac{\pi}{2} \end{pmatrix} + \begin{pmatrix} 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u-2 \\ v - \frac{\pi}{2} \end{pmatrix}$$

$$= u\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + v\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \pi \\ 0 \\ 0 \end{pmatrix},$$

para todo u y v. Como los vectores

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

son linealmente independientes, deducimos que la imagen de A es un plano.

Por tanto, la superficie S tiene plano tangente en $(0, 2, \pi/2)$. Eliminando u y v de las ecuaciones

$$x = -2v + \pi$$
$$y = u$$
$$z = v,$$

obtenemos

$$x = -2z + \pi$$

que es la ecuación que define implícitamente al plano tangente a S en $(0, 2, \pi/2)$. Véase Fig. 18.

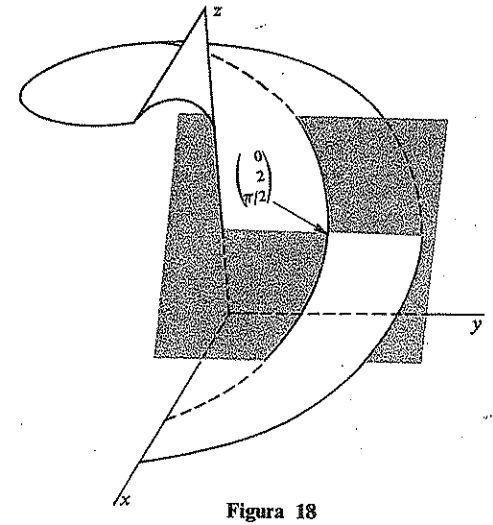


Figura 18

Ejemplo 2. La función de t definida por

$$f(t) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}, \qquad -\infty < t < \infty,$$

que ha sido discutida en el Ejemplo 6 de la Sec. 1 del Cap. 3, define paramétricamente la curva mostrada en la Fig. 19. La diferencial de f en t_0 está definida por la matriz jacobiana

$$\begin{pmatrix} 1 \\ 2t_0 \\ 3t_0^2 \end{pmatrix},$$

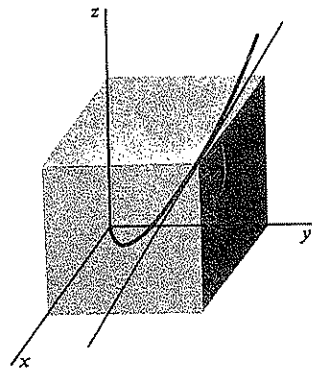


Figura 19

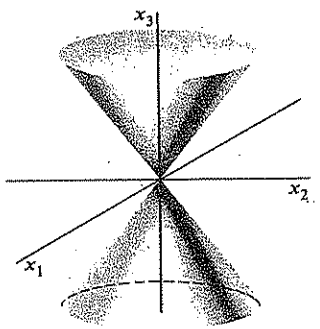


Figura 20

y la función afín

$$A(t) = f(t_0) + f'(t_0)(t - t_0)$$

que aproxima a f cerca de t_0, está dada por

$$A(t) = \begin{pmatrix} t_0 \\ t_0^2 \\ t_0^3 \end{pmatrix} + (t - t_0)\begin{pmatrix} 1 \\ 2t_0 \\ 3t_0^2 \end{pmatrix}$$

$$= t\begin{pmatrix} 1 \\ 2t_0 \\ 3t_0^2 \end{pmatrix} - \begin{pmatrix} 0 \\ t_0^2 \\ 2t_0^3 \end{pmatrix}, \qquad -\infty < t < \infty.$$

Como $(1, 2t_0, 3t_0^2) \neq 0$, tenemos que la imagen de A es la recta tangente a la curva en $f(t_0)$. La Fig. 19, muestra la curva y la recta tangente en $f(1) = (1, 1, 1)$.

La condición que la aproximación afín a $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ define un plano n-dimensional es importante, ya que impone restricciones a las tangentes y exige que la superficie sea suficientemente lisa si f es diferenciable. En cuanto a las tangentes, es claro que la dimensión del plano tangente en $f(\mathbf{x}_0)$ es igual a la de la imagen de la diferencial de f en $\mathbf{x}_0$. Pero como las columnas de la matriz jacobiana $f'(\mathbf{x}_0)$ (de forma m-por-n) generan la imagen de la diferencial, es suficiente exigir que esta matriz tenga n columnas linealmente independientes. Es decir,

7.1 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ diferenciable. Entonces el plano tangente a f en $f(\mathbf{x}_0)$ existe (y tiene dimensión n) si, y solo si, la matriz $f'(\mathbf{x}_0)$ tiene n columnas linealmente independientes. El requisito que f sea continuamente diferenciable significa que el plano tangente de una superficie lisa S varía continuamente de punto en punto de S. Para ver el efecto de la condición de dimensión, consideramos el ejemplo siguiente.

Ejemplo 3. La función $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}^3$ definida por

$$f(u, v) = \begin{pmatrix} u\cos v \\ u\,\mathrm{sen}\, v \\ u \end{pmatrix}, \qquad \text{para } (u, v) \text{ en } \mathcal{R}^2,$$

es continuamente diferenciable, ya que la matriz jacobiana

$$f'(u, v) = \begin{pmatrix} \cos v & -u\,\mathrm{sen}\, v \\ \mathrm{sen}\, v & u\cos v \\ 1 & 0 \end{pmatrix}$$

tiene elementos continuos. La imagen de f es el cono mostrado en la Fig. 20. Los puntos de este cono distintos del vértice, corresponden a $u \neq 0$ y es fácil comprobar que si $u \neq 0$, las columnas de $f'(u, v)$ son linealmente independientes. Por tanto, el plano tangente en esos puntos tiene la dimensión deseada, es decir, 2. Sin embargo, en el vértice, $u = 0$. Por tanto, $f'(0, v)$ tiene una sola columna distinta de cero. Por tanto, cualquier intento de usar la aproximación afín de f para definir una tangente al vértice conduce a una tangente 1-dimensional. En realidad, parece natural decir que el cono no tiene tangente en su vértice. Sin embargo, el cono satisface la definición de superficie lisa en todos los otros puntos. La falla en el vértice no está asociada con una falta de diferenciabilidad de f, sino que se debe a que la tangente no existe.

Superficies definidas explícitamente

Supongamos que el conjunto S está definido explícitamente por una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$. Esto significa que S es el gráfico de f en $\mathcal{R}^{n+m}$, formado por todos los puntos de la forma $(\mathbf{x}, f(\mathbf{x}))$, para cada $\mathbf{x}$ en el dominio de f. Si f es diferenciable en $\mathbf{x}_0$, la función afín

$$A(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

define explícitamente un plano n-dimensional y este plano se llama plano tangente de S en $\mathbf{x}_0$. Si, además, f es continuamente diferenciable, entonces S se llama una superficie lisa.

En este caso no existe la complicación de verificar que la diferencial tenga

imagen n-dimensional, como en el caso de la teoría paramétrica. Si $f'(\mathbf{x}_0)$ existe, entonces A tiene un gráfico que es siempre un plano n-dimensional.

7.2 Teorema

El gráfico de una función afín $\mathcal{R}^n \xrightarrow{A} \mathcal{R}^m$ es un plano n-dimensional.

Demostración. Por definición de función afín, existe una función lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^m$ y un vector $\mathbf{y}_0$ en $\mathcal{R}^m$ tales que

$$A(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0, \qquad \text{para todo } \mathbf{x} \text{ en } \mathcal{R}^n.$$

El gráfico de A es el conjunto de todos los puntos

$$(\mathbf{x}, A(\mathbf{x})) = (\mathbf{x}, L(\mathbf{x})) + (0, \mathbf{y}_0), \qquad \mathbf{x} \text{ en } \mathcal{R}^n.$$

Por tanto, ésta es la imagen bajo la traslación por $(0, \mathbf{y}_0)$ del gráfico de L. Por tanto, el gráfico de A es un plano n-dimensional si, y solo si, el gráfico de L, que es un subespacio de $\mathcal{R}^{n+m}$, tiene dimensión n. Pero si $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ es la base natural de $\mathcal{R}^n$, entonces todo punto $(\mathbf{x}, L(\mathbf{x}))$ en el gráfico de L puede ser escrito en la forma

$$(\mathbf{x}, L(\mathbf{x})) = x_1(\mathbf{e}_1, L(\mathbf{e}_1)) + \ldots + x_n(\mathbf{e}_n, L(\mathbf{e}_n)).$$

Es claro que los n vectores $(\mathbf{e}_j, L(\mathbf{e}_j))$ son linealmente independientes y, puesto que ellos generan el gráfico de L, éste tiene dimensión n.

Ejemplo 4 El hemisferio mostrado en la Fig. 21 está definido explícitamente por

$$g\begin{pmatrix} x \\ y \end{pmatrix} = \sqrt{9 - x^2 - y^2}.$$

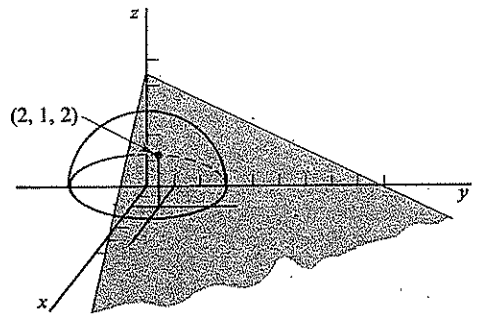


Figura 21

La diferencial de g en $\mathbf{x}_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ está definida por la matriz jacobiana

$$\left(\frac{-x}{\sqrt{9 - x^2 - y^2}} \quad \frac{-y}{\sqrt{9 - x^2 - y^2}}\right)_{x=2,y=1} = \left(-1 \quad -\frac{1}{2}\right).$$

El plano tangente al hemisferio en (2, 1, 2) es el gráfico de la función afín

$$\begin{aligned} A(\mathbf{x}) &= g(\mathbf{x}_0) + g'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= 2 + \left(-1 \quad -\frac{1}{2}\right)\begin{pmatrix} x - 2 \\ y - 1 \end{pmatrix} \\ &= \frac{9}{2} - x - \frac{1}{2}y. \end{aligned}$$

Este plano está definido implícitamente por la ecuación $z = A(\mathbf{x})$, es decir, por

$$2x + y + 2z = 9.$$

El gráfico de la función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ puede ser representado paramétricamente por una función $\mathcal{R}^n \xrightarrow{g} \mathcal{R}^{n+m}$ de la forma $g(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$. Si el conjunto S puede ser representado de ambas maneras, uno se pregunta si se obtienen las mismas tangentes y si la noción de superficie lisa es la misma en ambas representaciones. Es claro que g es continuamente diferenciable si, y solo si, f lo es. En efecto, $g'(\mathbf{x}_0)$ tiene la forma

$$g'(\mathbf{x}_0) = \begin{pmatrix} I \\ f'(\mathbf{x}_0) \end{pmatrix},$$

donde I es la matriz n-por-n identidad. Por tanto,

$$\begin{aligned} g(\mathbf{x}_0) + g'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) &= (\mathbf{x}_0, f(\mathbf{x}_0)) + (\mathbf{x} - \mathbf{x}_0, f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)) \\ &= (\mathbf{x}, f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)), \end{aligned}$$

lo cual muestra que el gráfico de la aproximación afín de f es igual a la imagen de la aproximación afín de f. Por tanto, en los casos comunes, las nociones de tangente y superficie lisa coinciden.

Recordamos que el núcleo de una función lineal $\mathcal{R}^{n+m} \xrightarrow{L} \mathcal{R}^m$ es un subespacio de $\mathcal{R}^{n+m}$. Por tanto, si el vector $\mathbf{x}_0$ es fijo, el conjunto de todos los $\mathbf{x}$ tales que $\mathbf{x} - \mathbf{x}_0$ están en el núcleo de L es un plano en $\mathcal{R}^{n+m}$. Claramente, este plano pasa por $\mathbf{x}_0$.

Funciones definidas implícitamente

Consideremos una función $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ y un vector fijo $\mathbf{z}_0$ en $\mathcal{R}^m$. Sea S el conjunto de nivel definido por la ecuación $F(\mathbf{x}) = \mathbf{z}_0$. Si F es diferenciable en

el punto $\mathbf{x}_0$ de $\mathcal{R}^{n+m}$ y la aproximación afín $A(\mathbf{x}) = F(\mathbf{x}_0) + F'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ determina implícitamente un plano n-dimensional por $A(\mathbf{x}) = \mathbf{z}_0$, entonces este plano se llama el tangente de S en $\mathbf{x}_0$. Como $F(\mathbf{x}_0) = \mathbf{z}_0$, la ecuación que define al plano se reduce a

$$F'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0. \tag{1}$$

Si, además, F es continuamente diferenciable en su dominio, entonces S es representable como una superficie lisa cerca de todos los puntos en los cuales hay un plano tangente.

Ejemplo 5. La ecuación $x^2 + y^2 + z^2 = 9$ define implícitamente una esfera en el espacio euclidiano tridimensional con radio 3 y centro en el origen. Una ecuación para el plano tangente a la esfera en

$$\mathbf{x}_0 = (x_0, y_0, z_0) = (2, 1, 2)$$

queda determinado como sigue: Si $F(x, y, z) = x^2 + y^2 + z^2$, entonces la matriz jacobiana $F'(\mathbf{x}_0)$ es

$$(2x_0 \quad 2y_0 \quad 2z_0) = (4 \quad 2 \quad 4).$$

La Ec. (1) que define implícitamente el plano tangente es, por tanto,

$$(4 \quad 2 \quad 4)\begin{pmatrix} x-2 \\ y-1 \\ z-2 \end{pmatrix} = 0.$$

Esta es equivalente a $4x + 2y + 4z = 18$, es decir, $2x + y + 2z = 9$. Nótese que hemos encontrado la misma ecuación que la obtenida en el Ejemplo 4. Si el plano determinado por la ecuación $F'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0$ no es de dimensión n, entonces, de acuerdo con la definición, la función $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ no le asigna un plano tangente al conjunto de nivel $F(\mathbf{x}) = \mathbf{z}_0$ en $\mathbf{x} = \mathbf{x}_0$. Esto es parecido a la complicación que aparece en la teoría paramétrica y que dio origen al Teorema 7.1. En el caso presente, necesitamos saber que el núcleo de la transformación lineal con matriz $F'(\mathbf{x}_0)$ tiene dimensión n. Como la dimensión del núcleo es igual a la dimensión del dominio menos la dimensión de la imagen (Teorema 4.7 del Cap. 2), lo que queremos es que la dimensión de la imagen sea m. Por tanto,

7.3 Teorema

Sea $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ diferenciable. Entonces el tangente al conjunto de nivel $F(\mathbf{x}) = \mathbf{z}_0$ en $\mathbf{x}_0$ existe (y tiene dimensión n) si, y solo si, $F'(\mathbf{x}_0)$ tiene m columnas linealmente independientes.

El Ejemplo 3 muestra que, en el caso paramétrico, una superficie puede no ser lisa debido a que la dimensión del tangente es demasiado pequeña. El ejemplo siguiente es bastante típico de la forma en que puede fallar el método estándar para definir tangentes.

Ejemplo 6. La Fig. 22 muestra el folio de Descartes, que es el conjunto de nivel determinado por $x^3 + y^3 - 3xy = 0$. La función $F(x, y) = x^3 + y^3 - 3xy$ es continuamente diferenciable con derivada

$$F'(x, y) = (3x^2 - 3y, 3y^2 - 3x).$$

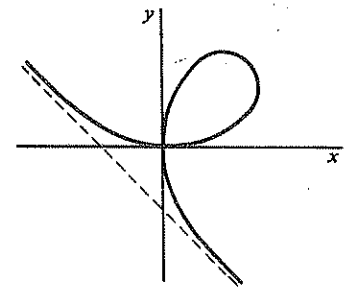


Figura 22

El criterio del Teorema 7.3 requiere que uno de los dos elementos sea distinto de cero, en cuyo caso, el tangente al punto (x_0, y_0) satisface

$$(x_0^2 - y_0)(x - x_0) + (y_0^2 - x_0)(y - y_0) = 0.$$

El único punto en que la diferencial no asocia un tangente es $(x, y) = (0, 0)$. En este punto el núcleo de la diferencial tiene dimensión 2.

La función $\mathcal{R} \xrightarrow{f} \mathcal{R}^2$ definida por

$$f(t) = \left(\frac{3t}{1+t^3}, \frac{3t^2}{1+t^3}\right), \qquad -1 < t < \infty$$

es una parametrización de la parte de la curva que está en el primer y segundo cuadrantes y ella asigna un tangente horizontal a la curva en el origen. Si intercambiamos las funciones coordenadas de f da una parametrización de la parte de la curva en el primer y cuarto cuadrantes.

$\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ es continuamente diferenciable y su conjunto de nivel $F(\mathbf{x}) = 0$ tiene un tangente en $\mathbf{x}_0$ determinado implícitamente; entonces, de acuerdo con el Teorema 7.3, la matriz m-por-$(n + m)$, $F'(\mathbf{x}_0)$, tiene m columnas linealmente independientes. Denotemos las variables que corresponden a estas co-

lumnas por el vector $\mathbf{y}$ y observemos que el teorema de la función implícita puede ser aplicado. De esto concluimos que si escribimos $\mathbf{x} = (\mathbf{v}, \mathbf{y})$, entonces existe una función continuamente diferenciable $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$ que satisface $F(\mathbf{v}, f(\mathbf{v})) = 0$ en una vecindad de $\mathbf{x}_0$. Este resultado dice que, en una vecindad de $\mathbf{x}_0$, el conjunto de nivel $F(\mathbf{x}) = 0$ puede ser representado como el gráfico de la función f. En el Ejercicio 12 dejamos la comprobación rutinaria de que se obtiene el mismo tangente con las representaciones implícitas y explícitas.

EJERCICIOS

1. Encuentre una representación paramétrica $t\mathbf{x}_1 + \mathbf{x}_2$ para la recta tangente a cada una de las curvas definidas paramétricamente por las siguientes funciones en los puntos indicados. Dibuje la curva y la recta tangente

(a) $f(t) = \begin{pmatrix} t \\ e^t \end{pmatrix}$, a $f(0)$. $\left[\textit{Resp}: t\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \right]$

(b) $g(t) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t^2 + 1 \\ t - 1 \\ t^2 \end{pmatrix}$, a $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$.

2. Encuentre el plano tangente a cada una de las superficies definidas paramétricamente por las siguientes funciones en los puntos indicados. Dibuje la superficie y el plano tangente en (a) y (b).

(a) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ u - v \\ u^2 - v^2 \end{pmatrix}$, $\begin{pmatrix} 0 \le u \le 2 \\ 0 \le v \le 2 \end{pmatrix}$ a $f\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$\left[\textit{Resp}: \begin{pmatrix} u + v + 2 \\ u - v \\ 2u - 2v \end{pmatrix}. \right]$

(b) $g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos u \operatorname{sen} v \\ \operatorname{sen} u \operatorname{sen} v \\ \cos v \end{pmatrix}$, $\begin{pmatrix} 0 \le u \le 2\pi \\ 0 \le v \le \pi/2 \end{pmatrix}$ a $g\begin{pmatrix} \pi \\ \pi/4 \end{pmatrix}$.

(c) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u + v + 1 \\ 2u + 3v \quad 1 \\ u + 2v - 2 \end{pmatrix}$, a $\begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}$.

3. Encuentre el plano o recta tangente en los puntos indicados, a las siguientes curvas o superficies definidas explícitamente.

(a) $f(x) = (x - 1)(x - 2)(x - 3)$, a $(0, -6)$.

(b) $f(x, y) = \dfrac{1}{x^2 + y^2}$, a $(x_0, y_0, f(x_0, y_0)) = (0, 2, \frac{1}{4})$.

(c) $g(t) = \begin{pmatrix} t \\ e^t \end{pmatrix}$, a $\begin{pmatrix} 1 \\ 1 \\ e \end{pmatrix}$. $\left[\textit{Resp}: L(t) = t\begin{pmatrix} 1 \\ e \end{pmatrix}. \right]$

(d) $g(x, y) = \cosh(x^2 + y^2)$, a $(x_0, y_0, g(x_0, y_0)) = (1, 2, \cosh 5)$.

4. Encuentre la recta o plano tangente en los puntos indicados, a las siguientes curvas o superficies definidas implícitamente. Dibuje la curva o superficie y su tangente en los casos (b), (d) y (e).

(a) $xy + yz + zx = 1$, a $\mathbf{x}_0 = (2, -1, 3)$.

(b) $\dfrac{x^2}{4} + \dfrac{y^2}{9} + z^2 = 1$, a $\mathbf{x}_0 = \left(1, 0, \dfrac{\sqrt{3}}{2}\right)$.

$\left[\textit{Resp}: \dfrac{x}{2} + \sqrt{3}\, z = 2. \right]$

(c) $5x + 5y + 2z = 8$ a $(1, 1, -1)$.

(d) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, a $(x_0, y_0) = \left(\dfrac{\sqrt{3}\, a}{2}, \dfrac{b}{2}\right)$.

(e) $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2 + y^2 + z^2 \\ x + y \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix}$, a $\mathbf{x}_0 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$.

$\left[\textit{Resp}: \begin{pmatrix} 4 & 2 & 4 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 18 \\ 3 \end{pmatrix}. \right]$

5. En cada uno de los Ejercicios 1(a), 2(a), 2(b) y 4(b) encuentre una normal a la curva o superficie en el punto indicado.

$\left[\textit{Resp}: 1(a) \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \right]$

6. De acuerdo con nuestra definición, cada una de las curvas o superficies siguientes, no tiene una recta o plano tangente en el punto indicado. ¿Por qué?

(a) $f(t) = \begin{pmatrix} t \\ |t| \\ t^2 \end{pmatrix}$, a $f(0)$.

(b) $g(t) = \begin{pmatrix} t^2 - 2t \\ t^3 - 3t \\ t^4 - t^3 - t \end{pmatrix}$, a $\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}$.

(c) $f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2v^4 \\ uv^2 \\ u^2 + v^4 \end{pmatrix}$, a $f\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(d) $f(x, y) = \sqrt{1 - x^2 - y^2}$, a $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)$.

(e) $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z^2e^{x+y} \\ 2xyz^2 \end{pmatrix} = \begin{pmatrix} 4e^2 \\ 8 \end{pmatrix}$, a $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

7. Encuentre todos los puntos en los cuales la superficie definida paramétricamente por la función

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2v^2 \\ uv \\ uv + 1 \end{pmatrix}$$

no tiene un plano tangente.

8. Dos funciones vectoriales pueden definir la misma curva o superficie. Demuestre que las funciones

$$f_1(t) = (\cos t, \operatorname{sen} t), \qquad 0 < t < 2\pi,$$

$$f_2(s) = \left(\frac{s^2 - 1}{s^2 + 1}, \frac{2s}{s^2 + 1}\right), \quad -\infty < s < \infty,$$

definen paramétricamente la misma curva en el espacio euclidiano de dos dimensiones.

9. Considere las funciones vectoriales

$$f(t) = \begin{pmatrix} t^3 \\ t^3 \\ t^3 \end{pmatrix}, \qquad -\infty < t < \infty,$$

$$g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^3 \\ v^3 \\ 0 \end{pmatrix}, \qquad \begin{cases} -\infty < u < \infty, \\ -\infty < v < \infty. \end{cases}$$

(a) ¿Cuál curva y cuál superficie están definidas paramétricamente por f y g?
(b) Demuestre que, de acuerdo con nuestre definición, la curva no tiene una recta tangente y la superficie no tiene un plano tangente en (0, 0, 0).

10. Sea

$$f(t) = \begin{pmatrix} t^3 \\ t^3 \end{pmatrix}, \qquad -\infty < t < \infty.$$

(a) Demuestre que la curva en $\mathfrak{R}^3$ definida explícitamente por f tiene una recta tangente en cada uno de sus puntos.
(b) Demuestre que la curva en $\mathfrak{R}^2$ definida paramétricamente por f no tiene recta tangente en un punto.
(c) Interprete geométricamente (a) y (b). ¿Cuál es la relación entre la recta tangente de (a) y la de (b)?

11. Sea $\mathbf{y}_0 = 0$ un punto de la imagen de la función $\mathfrak{R}^n \xrightarrow{F} \mathfrak{R}^m$. Suponga que la superficie S definida implícitamente por la ecuación $F(\mathbf{x}) = 0$ tiene una tangente $\mathcal{T}$ en $\mathbf{x}_0$.
(a) Compruebe que la superficie S' en $\mathfrak{R}^{n+m}$ definida explícitamente por F tiene una tangente $\mathcal{T}'$ en $(\mathbf{x}_0, \mathbf{y}_0)$.
(b) Sea P el plano en $\mathfrak{R}^{n+m}$ definido por la ecuación $\mathbf{y} = 0$. Demuestre que $S = S' \cap P$.
(c) Demuestre que $\mathcal{T} = \mathcal{T}' \cap P$.
(d) Usando la ecuación $F(x, y) = x^2 + y^2 - 2 = 0$ y el punto $\mathbf{x}_0 = (1, 1)$ haga un dibujo que muestre a S, S', $\mathcal{T}$ y P. ¿Qué sucede con $\mathcal{T}'$?

12. Demuestre que si $\mathfrak{R}^{n+m} \xrightarrow{F} \mathfrak{R}^m$ define una superficie lisa S entonces su tangente en $\mathbf{x}_0$ determinada por $F'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0$ es igual a la tangente del gráfico de la función $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ que satisface $F(\mathbf{v}, f(\mathbf{v})) = 0$ y cuya existencia está garantizada por el teorema de la función implícita.

13. Verifique que las dos tangentes definidas paramétricamente en el origen descritas en el Ejemplo 6 son horizontal y vertical respectivamente.

14. Demuestre que si P es un plano n-dimensional que pasa por el origen en $\mathfrak{R}^{n+m}$ entonces P es el núcleo de una función lineal con valores en $\mathfrak{R}^m$.

5

Funciones reales

1. Valores extremos

El problema de encontrar los valores máximos y mínimos de una función real f de más variables es importante porque si se conoce el comportamiento extremo de f, entonces, podemos acotar sus valores posibles. Además, puede ser importante para cuáles elementos del dominio de f se alcanzan los valores extremos. Ejemplos familiares son los extremos de la temperatura, velocidad o ganancia económica, cada uno de los cuales puede ser, en un problema práctico, función de varias variables.

Una función real f tiene un **valor máximo absoluto en** $\mathbf{x}_0$ si, para todo $\mathbf{x}$ en el dominio de f,

$$f(\mathbf{x}) \leq f(\mathbf{x}_0),$$

y un **valor mínimo absoluto en** $\mathbf{x}_0$ si, por el contrario,

$$f(\mathbf{x}_0) \leq f(\mathbf{x}).$$

El número $f(\mathbf{x}_0)$ se llama un **valor máximo local** o **valor mínimo local** si existe una vecindad N de $\mathbf{x}_0$ tal que, respectivamente,

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{o} \quad f(\mathbf{x}_0) \leq f(\mathbf{x}),$$

para todo $\mathbf{x}$ en N. Un valor máximo o mínimo de f se llama un **valor extremo**. Un punto $\mathbf{x}_0$ en el cual f toma un valor extremo se llama un **punto extremo**.

Ejemplo 1. Consideremos la función $f(x, y) = x^2 + y^2$, cuyo dominio es el conjunto de todos los puntos (x, y) que están dentro de o en la elipse $x^2 + 2y^2 = 1$. La Fig. 1 muestra el gráfico de f. Supongamos que f tiene un valor extremo (es decir, un máximo o mínimo) en el punto (x_0, y_0) en el interior de la elipse. Entonces es evidente que las funciones f_1 y f_2 definidas por

$$f_1(x) = f(x, y_0), \qquad f_2(y) = f(x_0, y)$$

deben tener valores extremos en x_0 e y_0, respectivamente. Aplicando el criterio bien conocido para las funciones de una variable, tenemos que

$$f_1'(x_0) = f_2'(y_0) = 0.$$

Como

$$f_1'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0)$$

y

$$f_2'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0),$$

una condición necesaria para que f tenga un valor extremo en (x_0, y_0) es

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

En este ejemplo,

$$\frac{\partial f}{\partial x}(x, y) = 2x \quad \text{y} \quad \frac{\partial f}{\partial y}(x, y) = 2y,$$

de modo que el único valor extremo de f en el interior de la elipse está en el punto $(x_0, y_0) = 0$. Mirando el gráfico de f, dado en la Fig. 1, es claro que este valor es un mínimo. Consideramos ahora los valores de f en la curva frontera. La elipse puede ser definida paramétricamente por la función

$$g(t) = (x, y) = \left(\cos t, \frac{1}{\sqrt{2}} \operatorname{sen} t\right), \qquad 0 \leq t < 2\pi.$$

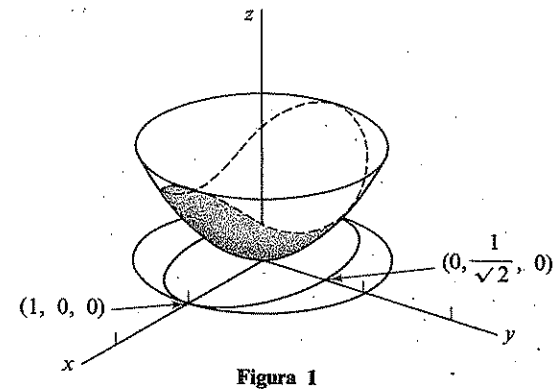


Figura 1

Por tanto, los valores de f en la elipse son los valores de la función compuesta $f \circ g$. Todo valor extremo de f en la elipse será extremo de $f \circ g$. Esta última es una función real de una variable y la tratamos de la forma habitual, es decir, hacemos su derivada igual a cero. Por la regla de la cadena, obtenemos

$$\frac{d}{dt}(f \circ g) = \left(2 \cos t \frac{2}{\sqrt{2}} \operatorname{sen} t\right)\begin{pmatrix} -\operatorname{sen} t \\ \frac{1}{\sqrt{2}} \cos t \end{pmatrix}$$

$$= -2 \cos t \operatorname{sen} t + \operatorname{sen} t \cos t$$

$$= -\tfrac{1}{2} \operatorname{sen} 2t.$$

Por tanto, puede haber valores extremos en $t = 0$, $\pi/2$, π y $3\pi/2$. Los valores correspondientes de (x, y) son $(1, 0)$, $(0, 1/\sqrt{2})$, $(-1, 0)$ y $(0, 1/\sqrt{2})$ y los de f son $1, \frac{1}{2}$, 1 y $\frac{1}{2}$, respectivamente. Vemos así que el mínimo absoluto de f es 0 en $(0, 0)$ y que el máximo absoluto de f ocurre en los puntos $(1, 0)$ y $(-1, 0)$. Nótese que los valores extremos de $f \circ g$ que ocurren en $t = \pi/2$ y $3\pi/2$ no son extremos de f, como puede comprobarse mirando la Fig. 1.

Los métodos usados en el ejemplo precedente son válidos en cualquier dimensión. El teorema siguiente es el criterio principal usado en esta generalización y, a pesar de que él puede ser demostrado reduciéndolo al caso de una variable, damos una demostración que contiene como caso especial la situación de una variable.

1.1 Teorema

Si la función diferenciable $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ tiene un valor extremo local en el punto $\mathbf{x}_0$, que está en el interior de su dominio, entonces $f'(\mathbf{x}_0) = 0$.

Demostración. Supongamos que f tiene un mínimo local en $\mathbf{x}_0$. Para todo vector unitario $\mathbf{u}$ en $\mathcal{R}^n$, existe $\epsilon > 0$ tal que si $-\epsilon < t < \epsilon$, entonces $f(\mathbf{x}_0) \leq f(\mathbf{x}_0 + t\mathbf{u})$. Por tanto, si $0 < t < \epsilon$, entonces

$$0 \leq \frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t},$$

$$0 \leq \frac{f(\mathbf{x}_0 - t\mathbf{u}) - f(\mathbf{x}_0)}{t}.$$

Por el Teorema 1.1 del Cap. 4, se tiene que

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = f'(\mathbf{x}_0)\mathbf{u}.$$

Por tanto,

$$0 \leq \lim_{t \to 0+} \frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t} = f'(\mathbf{x}_0)\mathbf{u},$$

$$0 \leq \lim_{t \to 0+} \frac{f(\mathbf{x}_0 - t\mathbf{u}) - f(\mathbf{x}_0)}{t} = f'(\mathbf{x}_0)(-\mathbf{u}) = -f'(\mathbf{x}_0)\mathbf{u}.$$

De lo cual concluimos que $f'(\mathbf{x}_0)\mathbf{u} = 0$. Como $\mathbf{u}$ es arbitrario, $f'(\mathbf{x}_0) = 0$. El argumento para un valor máximo es análogo.

El resultado indicado era de esperar. Recordemos que

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = f'(\mathbf{x}_0)\mathbf{u},$$

y que la derivada con respecto a $\mathbf{u}$ mide la velocidad de cambio de f en la dirección de $\mathbf{u}$. En un punto extremo, en el interior del dominio de f, esta velocidad debería ser cero en toda dirección. La importancia del teorema está en el hecho que los únicos puntos del interior del dominio de f que debemos considerar al buscar puntos extremos son aquellos para los cuales $f'(\mathbf{x}) = 0$. Los puntos $\mathbf{x}$ para los cuales $f'(\mathbf{x}) = 0$ se llaman **puntos críticos** de f.

En la práctica, a menudo tenemos una función f que es diferenciable en un conjunto abierto y queremos encontrar los puntos extremos de f cuando ella se restringe a un subconjunto S del dominio de f. En el ejemplo que sigue se ilustra las dos afirmaciones siguientes: (1) *un punto* $\mathbf{x}$ *tal que* $f'(\mathbf{x}) = 0$ *no es necesariamente un punto extremo de* f; (2) f *puede tener un punto extremo* $\mathbf{x}$ *cuando se restringe a un conjunto* S, *sin que* $f'(\mathbf{x}) = 0$.

Ejemplo 2. Sea $f(x, y, z) = xyz$ en la región definida por $|x| \leq 1$, $|y| \leq 1$, $|z| \leq 1$. Es decir, el dominio de f es el cubo de arista 2 mostrado en la Fig. 2. La condición $f'(\mathbf{x}) = 0$ para puntos críticos, es equivalente a $(yz, xz, xy) = (0, 0, 0)$. Las soluciones de esta ecuación son los puntos que satisfacen $x = y = 0$ o $x = z = 0$ o $y = z = 0$; es decir, los ejes coordenados. Como f tiene el valor cero en cada uno de sus puntos críticos y como f tiene valores positivos y negativos cerca de cada uno de esos puntos, ningún punto crítico es un punto extremo. Además, es claro que f tiene el valor máximo 1 y el valor mínimo -1. Estos valores aparecen en las aristas del cubo, ninguna de las cuales es un punto crítico.

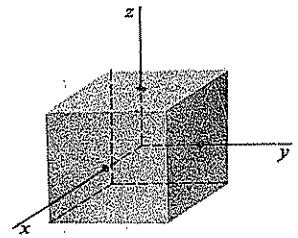


Figura 2

El problema de encontrar los valores extremos de una función f en la frontera de una subregión R de $\mathcal{R}^n$ es un caso especial del caso en que f ha sido restringida a un conjunto S de dimensión menor que la de R. Entonces, como viéramos en el Ejemplo 2, no es suficiente examinar los puntos críticos de f como función en R. Más generalmente, podemos estar interesados en f cuando se la restringe a un conjunto S de dimensión menor, que no necesariamente la frontera de una región.

Ejemplo 3. La función $f(x, y, z) = y^2 - z - x$ tiene una diferencial definida por la matriz

$$(-1 \quad 2y \quad -1),$$

de modo que f no tiene puntos críticos cuando se la considera como una función definida en $\mathfrak{R}^3$. Por otra parte, supongamos que restringimos f a la curva definida paramétricamente por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}, \qquad -\infty < t < \infty.$$

En γ, f toma los valores $F(t) = f(t, t^2, t^3) = t^4 - t^3 - t$ cuando t varía en $(-\infty, \infty)$. Tenemos que

$$F'(t) = 4t^3 - 3t^2 - 1 = (t - 1)(4t^2 + t + 1).$$

Por tanto, $F'(t)$ es cero solo en $t = 1$. Además, como $F''(t) = 12t^2 - 6t$, tenemos que $F''(1) > 0$. Por tanto, el punto $(1, 1, 1)$ es un mínimo relativo de f restringida a la curva γ. El valor mínimo de f en γ es -1 y no hay otros valores extremos.

Ejemplo 4. Supongamos que la función $f(x, y, z) = x + y + z$ se restringe a la intersección de las dos superficies

$$x^2 + y^2 = 1, \qquad z = 2$$

mostradas en la Fig. 3. La curva de intersección C puede ser parametrizada por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos t \\ \operatorname{sen} t \\ 2 \end{pmatrix}, \qquad 0 \leq t < 2\pi.$$

En C, la función f toma los valores $F(t) = \cos t + \operatorname{sen} t + 2$. Tenemos que $F'(t) = -\operatorname{sen} t + \cos t$, de modo que $F'(t) = 0$ en $t = \pi/4$ y $t = 5\pi/4$. Como $F''(\pi/4) < 0$ y $F''(5\pi/4) > 0$,

$$f(\sqrt{2}/2, \sqrt{2}/2, 2) = \sqrt{2} + 2$$

es el valor máximo y

$$f(-\sqrt{2}/2, -\sqrt{2}/2, 2) = -\sqrt{2} + 2$$

es el valor mínimo de f en C.

La solución del problema anterior dependió de la posibilidad de encontrar una representación paramétrica correcta de la curva formada por la intersección del cilindro $x^2 + y^2 - 1 = 0$ con el plano $z - 2 = 0$. Cuando no se dispone de una parametrización, a veces podemos utilizar el método de los **multiplicadores de Lagrange** que procederemos a discutir. Este método consiste en la verificación de la existencia de una representación paramétrica y en la deducción de

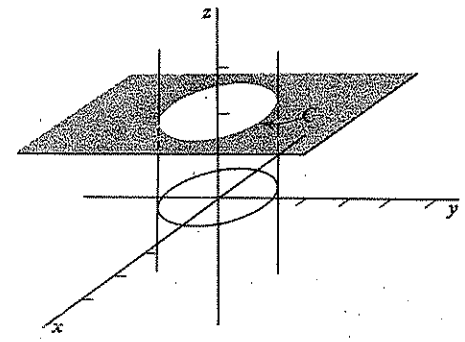


Figura 3

condiciones necesarias para la existencia de un punto extremo de una función f cuando se la restringe a la superficie o curva parametrizada.

1.2 Teorema. Método de los multiplicadores de Lagrange

Supongamos que la función $\mathfrak{R}^n \xrightarrow{G} \mathfrak{R}^m$, $n > m$, es continuamente diferenciable y que sus funciones coordenadas son $G_1, \ldots, G_m$.

Supongamos que las ecuaciones

$$\begin{aligned} G_1(x_1, \ldots, x_n) &= 0 \\ G_2(x_1, \ldots, x_n) &= 0 \\ &\cdots \\ G_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

definen implícitamente una superficie S en $\mathfrak{R}^n$ y que en el punto $\mathbf{x}_0$ de S la matriz $G'(\mathbf{x}_0)$ tiene m columnas linealmente independientes.

Si $\mathbf{x}_0$ es un punto extremo de una función diferenciable $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}$ restringida a S, entonces $\mathbf{x}_0$ es un punto crítico de la función

$$f + \lambda_1 G_1 + \ldots + \lambda_m G_m$$

para algunas constantes $\lambda_1, \ldots, \lambda_m$.

La demostración de este teorema será dada en el Apéndice. Damos aquí un argumento para hacer plausible este resultado. Recordemos que el gradiente de una función real diferenciable definida en $\mathfrak{R}^n$ es la función vectorial ∇f definida por

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right).$$

En términos del gradiente, la condición de Lagrange

$$f'(\mathbf{x}_0) + \lambda_1 G_1'(\mathbf{x}_0) + \ldots + \lambda_m G_m'(\mathbf{x}_0) = 0$$

puede ser expresada como

$$\nabla f(\mathbf{x}_0) + \lambda_1 \nabla G_1(\mathbf{x}_0) + \ldots + \lambda_m \nabla G_m(\mathbf{x}_0) = 0. \tag{1}$$

La Ec. 2.4 del Cap. 4 dice que el vector $\nabla G_i(\mathbf{x}_0)$ es perpendicular en $\mathbf{x}_0$ a la superficie S_i definida implícitamente por $G_i(\mathbf{x}) = 0$. Pero enconces cada vector $\nabla G_i(\mathbf{x}_0)$ es perpendicular en el mismo punto a la intersección S de todas las superficies S_i. Como por la Ec. (1) $\nabla f(\mathbf{x}_0)$ es una combinación lineal de los vectores $\nabla G_i(\mathbf{x}_0)$, el gradiente de f es perpendicular a S en $\mathbf{x}_0$. Ahora recordemos que, por el Teorema 2.2 del Cap. 4, el gradiente de f tiene la dirección de aumento máximo de f. Es razonable que esta dirección sea perpendicular a S en un punto extremo $\mathbf{x}_0$ de f restringida a S. En efecto, en caso contrario podríamos esperar encontrar un valor mayor o menor de f moviéndonos a lo largo en la dirección de $\nabla f(\mathbf{x}_0)$ proyectada en S.

Al aplicar este teorema es importante verificar que $G'(\mathbf{x})$ tiene m columnas independientes para $\mathbf{x}$ en S. Los puntos en que esta condición falla deben ser examinados por separado. Todos los puntos extremos $\mathbf{x}_0$, para los cuales la condición es satisfecha, son tales que existen constantes $\lambda_1, \ldots, \lambda_m$ para las cuales

$$f + \lambda_1 G_1 + \ldots + \lambda_m G_m$$

tiene a $\mathbf{x}_0$ como un punto crítico; es decir,

$$f'(\mathbf{x}_0) + \lambda_1 G_1'(\mathbf{x}_0) + \ldots + \lambda_m G_m'(\mathbf{x}_0) = 0. \tag{2}$$

Ejemplo 5. El problema del Ejemplo 4 consiste en encontrar los puntos extremos de $f(x, y, z) = x + y + z$ bajo las condiciones

$$x^2 + y^2 - 1 = 0, \qquad z - 2 = 0. \tag{3}$$

Escribimos

$$(x + y + z) + \lambda_1(x^2 + y^2 - 1) + \lambda_2(z - 2).$$

Los puntos críticos de esta función ocurren cuando

$$1 + 2\lambda_1 x = 0, \qquad 1 + 2\lambda_1 y = 0, \qquad 1 + \lambda_2 = 0.$$

Además, las Ecs. (3) deben ser satisfechas. Despejando λ_1, λ_2, x, y y z, tenemos

$$\lambda_2 = -1, \qquad \lambda_1 = \pm\frac{1}{\sqrt{2}}, \qquad x = y = \pm\frac{1}{\sqrt{2}}, \qquad z = 2.$$

Por tanto, los puntos críticos son

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 2\right) \quad \text{y} \quad \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 2\right).$$

Igual que en el Ejemplo 4, vemos fácilmente que f tiene su valor máximo, $\sqrt{2} + 2$, en el primero de estos puntos y su valor mínimo, $-\sqrt{2} + 2$, en el otro. Nótese que, a pesar de que no se necesitan todos los λ para la respuesta final, es necesario considerar todos los valores de los λ para los cuales se pueden satisfacer las ecuaciones.

Ejemplo 6. Encontremos el valor máximo de $f(x, y, z) = x - y + z$, sujeto a las condiciones $x^2 + y^2 + z^2 = 1$. La función

$$x - y + z + \lambda(x^2 + y^2 + z^2 - 1)$$

tiene puntos críticos que satisfacen

$$1 + 2\lambda x = 0, \qquad -1 + 2\lambda y = 0, \qquad 1 + 2\lambda z = 0,$$

y

$$x^2 + y^2 + z^2 = 1.$$

Las soluciones de estas ecuaciones son

$$\lambda = \pm\frac{3}{\sqrt{2}}, \qquad x = -y = z = \pm\frac{1}{\sqrt{3}}.$$

El máximo de f ocurre en $(1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3})$. El valor máximo es $\sqrt{3}$.

Ejemplo 7. Supongamos que $g(x_1, x_2, \ldots, x_n) = 0$ define implícitamente una superficie S en $\mathcal{R}^n$ y sea $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ un punto fijo. Minimizar la distancia de S a $\mathbf{a}$ es lo mismo que minimizar el cuadrado de la distancia. Por tanto, el punto $\mathbf{p}$ en S más cercano a $\mathbf{a}$ debe estar entre los puntos críticos de

$$\sum_{k=1}^{n}(x_k - a_k)^2 + \lambda g(x_1, \ldots, x_n)$$

para algún λ. Además de $g(x_1, \ldots, x_n) = 0$, los puntos críticos satisfacen las ecuaciones

$$2(x_1 - a_1) + \lambda\frac{\partial g}{\partial x_1}(x_1, \ldots, x_n) = 0$$

$$\vdots$$

$$2(x_n - a_n) + \lambda\frac{\partial g}{\partial x_n}(x_1, \ldots, x_n) = 0.$$

En el punto crítico $\mathbf{p}$ estas ecuaciones tienen la forma vectorial

$$\begin{pmatrix} p_1 - a_1 \\ \vdots \\ p_n - a_n \end{pmatrix} = -\frac{\lambda}{2}\begin{pmatrix} \dfrac{\partial g}{\partial x_1}(\mathbf{p}) \\ \vdots \\ \dfrac{\partial g}{\partial x_n}(\mathbf{p}) \end{pmatrix},$$

donde $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$. Por tanto, el vector $\mathbf{p} - \mathbf{a}$ de la izquierda es paralelo al vector normal a S en $\mathbf{p}$. Este vector normal aparece en el término de la derecha en la ecuación. En otras palabras, $\mathbf{p} - \mathbf{a}$ es perpendicular a S. Un ejemplo en dos dimensiones se ilustra en la Fig. 4.

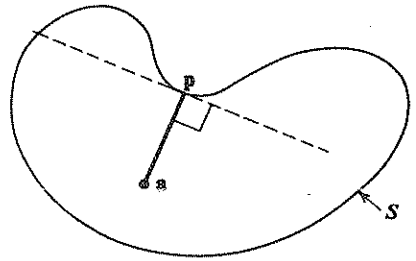


Figura 4

Ejemplo 8. Supongamos que un tubo cilíndrico debe contener un volumen fijo V y que su superficie, incluyendo la tapa y el fondo debe ser lo más pequeña posible. Si el radio del tubo es x y su altura es y, entonces $V = \pi x^2 y$. Queremos minimizar el área total $2\pi x^2 + 2\pi xy$. Escribimos

$$f(x, y) = 2\pi x^2 + 2\pi xy + \lambda(\pi x^2 y - V)$$

y buscamos los puntos críticos de f. Tenemos que $f_x = 0$, $f_y = 0$ se reduce a

$$2x + y + \lambda xy = 0, \qquad 2x + \lambda x^2 = 0.$$

La segunda ecuación es satisfecha por $x = 0$ o por $\lambda x = -2$. Pero si $x = 0$, entonces $V = 0$. Por tanto, sustituimos $\lambda x = -2$ en la primera ecuación, obteniendo $2x = y$. Por tanto, la altura y debe ser igual al diámetro $2x$. El valor de x para un volumen dado V se obtiene de la ecuación $2\pi x^3 = \pi x^2 y = V$.

Ejemplo 9. Los planos

$$x + y + z - 1 = 0 \quad \text{y} \quad x + y - z = 0$$

se intersectan en la recta S mostrada en la Fig. 5. Sea $f(x, y) = xy$ y restrinjamos f a la recta S. Usando el método de Lagrange, consideramos

$$xy + \lambda(x + y + z - 1) + \mu(x + y - z).$$

Sus puntos críticos ocurren cuando

$$y + \lambda + \mu = 0, \qquad x + \lambda + \mu = 0, \qquad \lambda - \mu = 0.$$

El único punto que satisface estas condiciones y que además está en S es $\mathbf{x}_0 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$. Tenemos que $\nabla f(\mathbf{x}_0) = (\frac{1}{4}, \frac{1}{4}, 0)$, que es perpendicular a S. La Fig. 5 muestra al vector unitario $\mathbf{u}$ en la dirección de $\nabla f(\mathbf{x}_0)$ con su punto inicial en $\mathbf{x}_0$.

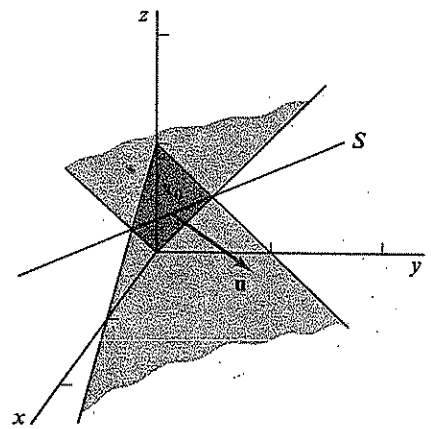


Figura 5

EJERCICIOS

1. Encuentre los puntos críticos de $x^2 + 4xy - y^2 - 8x - 6y$. [*Resp.* (2, 1).]
2. Encuentre los puntos en los cuales las funciones siguientes alcanzan sus valores más grandes y más pequeños.
   (a) $x + y$ en el cuadrado de vértices $(\pm 1, \pm 1)$.
   [*Resp.* Max (1, 1), min $(-1, -1)$.]
   (b) $x + y + z$ en la región $x^2 + y^2 + z^2 \leq 1$.
   [*Resp.* Max $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, min $(-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3})$.]
   (c) $x^2 + 24xy + 8y^2$ en la región $x^2 + y^2 \leq 25$.
   [*Resp.* Max $\pm(3, 4)$, min $\pm(4, -3)$.]
   (d) $1/(x^2 + y^2)$ en la región $(x - 2)^2 + y^2 \leq 1$.
   [*Resp.* Max (1, 0), min (3, 0).]
   (e) $x^2 + y^2 + (2\sqrt{2}/3)xy$ en la elipse $x^2 + 2y^2 \leq 1$.
   [*Resp.* Max $(\pm 2/\sqrt{5}, \pm 1/\sqrt{10})$, min (0, 0).]
3. Encuentre los puntos más lejanos del origen que están en la curva
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos t \\ \operatorname{sen} t \\ \operatorname{sen} t/2 \end{pmatrix}$$
   [*Resp.* $-1, -1$).]
4. Encuentre los puntos críticos de las siguientes funciones:
   (a) $x + y \operatorname{sen} x$. (b) $xy + xz$. (c) $x^2 + y^2 + z^2 - 1$.
5. Encuentre el valor máximo de la función $x(y + z)$ dado que $x^2 + y^2 = 1$ y $xz = 1$. [*Resp.* $\frac{3}{2}$.]

6. Encuentre el valor mínimo de $x + y^2$ con la condición $2x^2 + y^2 = 1$.
[*Resp.* $-1/\sqrt{2}$.]

7. Sean $f(x, y)$ y $g(x, y)$ continuamente diferenciables y suponga que, bajo la condición $g(x, y) = 0$, $f(x, y)$ alcanza su valor máximo M en (x_0, y_0). Demuestre que la curva de nivel $f(x, y) = M$ es tangente a la curva $g(x, y)$ en (x_0, y_0).

8. Una caja rectangular sin tapa debe tener una superficie con un área de 32 unidades. Encuentre las dimensiones que le darán un volumen máximo.

9. Encuentre la distancia mínima en $\mathcal{R}^2$ entre la elipse $x^2 + 2y^2 = 1$ y la recta $x + y = 4$. [*Indicación*: Considere el cuadrado de la distancia como una función de cuatro variables.]

10. (a) Encuentre el valor máximo de $x^2 + xy + y^2 + yz + z^2$ con la condición $x^2 + y^2 + z^2 = 1$. [*Resp.* $1 + 1/\sqrt{2}$.]
(b) Encuentre el valor mínimo de la misma función con las condiciones $x^2 + y^2 = 1$ y $ax + by + cz = 0$, donde (a, b, c) es el punto en el cual se alcanza el máximo obtenido en (a). [*Resp.* 1.]

11. Considere una función diferenciable $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ y una función continuamente diferenciable $\mathcal{R}^n \xrightarrow{G} \mathcal{R}^m$, $m < n$. Suponga que la superficie S definida por $G(x) = 0$ tiene una tangente $\mathcal{T}$ de dimensión $n - m$ en $\mathbf{x}_0$ y que la función f restringida a S tiene un valor extremo en $\mathbf{x}_0$. Demuestre que $\mathcal{T}$ es paralelo al tangente a la superficie definida explícitamente por f, en el punto $(\mathbf{x}_0, f(\mathbf{x}_0))$.

12. (a) Encuentre los puntos $\mathbf{x}_0$ en los cuales $f(x, y) = x^2 - y^2 - y$ alcanza su máximo en el círculo $x^2 + y^2 = 1$. [*Resp.* $(\pm\sqrt{15}/5, -\frac{1}{4})$.]
(b) Encuentre las direcciones en que f aumenta más rápidamente en $\mathbf{x}_0$. [*Resp.* $(\pm\sqrt{15}/4, -\frac{1}{4})$.]

13. Los planos $x + y - z - 2w = 1$ y $x - y + z + w = 2$ se cortan en un conjunto $\mathcal{F}$ en $\mathcal{R}^4$. Encuentre el punto de $\mathcal{F}$ más cercano al origen. [*Resp.* $(\frac{27}{19}, -\frac{7}{19}, \frac{7}{19}, -\frac{3}{19})$.]

14. Sean $\mathbf{x}_1, \ldots, \mathbf{x}_N$ puntos en $\mathcal{R}^n$ y sea

$$f(\mathbf{x}) = \sum_{k=1}^{N} |\mathbf{x} - \mathbf{x}_k|^2.$$

Encuentre el punto en el cual f alcanza su máximo y calcule el valor mínimo de f.

15. Demuestre, resolviendo un problema adecuado de mínimo, que si $a_k > 0$, $k = 1, \ldots, n$, entonces

$$(a_1 a_2 \ldots a_n)^{1/n} \leq \frac{a_1 + a_2 + \ldots + a_n}{n}$$

2. Polinomios cuadráticos

Sea $F(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ el producto punto euclidiano de dos vectores $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$. Además de tener la propiedad que $F(\mathbf{x}, \mathbf{x}) \geq 0$, la función F satisface

$$F(\mathbf{x}, \mathbf{y}) = F(\mathbf{y}, \mathbf{x}) \tag{1}$$

$$F(\mathbf{x} + \mathbf{x}', \mathbf{y}) = F(\mathbf{x}, \mathbf{y}) + F(\mathbf{x}', \mathbf{y}) \tag{2}$$

$$F(a\mathbf{x}, \mathbf{y}) = aF(\mathbf{x}, \mathbf{y}), \tag{3}$$

donde a es cualquier número real. Como consecuencia de la definición de F o de las propiedades antes indicadas, F también es lineal en la segunda variable. Debido a la propiedad de simetría (1) y a la linealidad en las dos variables, una función real F que satisface (1)-(3) para todo par de vectores $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$, se llama una **función bilineal simétrica**. Una tal función puede ser escrita en términos de coordenadas de la manera siguiente. Sea $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$. Entonces

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i, \qquad \mathbf{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j,$$

donde $\mathbf{e}_k$, $k = 1, 2, \ldots, n$ son los vectores de la base natural

$$\begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix},$$

de $\mathcal{R}^n$. De (2) y (3) se deduce que

$$F(\mathbf{x}, \mathbf{y}) = F\left(\sum_1^n x_i \mathbf{e}_i, \sum_1^n y_j \mathbf{e}_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} F(\mathbf{e}_i, \mathbf{e}_j) x_i y_j,$$

donde, por (1), $F(\mathbf{e}_i, \mathbf{e}_j) = F(\mathbf{e}_j, \mathbf{e}_i)$. Conversamente, una elección arbitraria de los números $F(\mathbf{e}_i, \mathbf{e}_j) = a_{ij}$, con la condición $a_{ij} = a_{ji}$, determina la función bilineal simétrica más general. Resumiendo, las funciones bilineales simétricas son aquellas que, en términos de sus coordenadas, tienen la forma

$$F(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} a_{ij} x_i y_j, \qquad a_{ij} = a_{ji}. \tag{4}$$

En particular, si $a_{ij} = 1$ y $a_{ij} = 0$ para $i \neq j$, obtenemos nuestro ejemplo original

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Si F es una función bilineal simétrica en $\mathcal{R}^n$, la función real de un vector, definida por $Q(\mathbf{x}) = F(\mathbf{x}, \mathbf{x})$. para todo $\mathbf{x}$ en $\mathcal{R}^n$, se llama un **polinomio cuadrático homogéneo** o, a veces, una **forma cuadrática**. Entonces, por definición, toda Q está asociada a una F bilineal y viceversa. De (4) se deduce que, en coordenadas,

$$Q(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij} x_i y_j, \qquad a_{ij} = a_{ji}. \tag{5}$$

Por ejemplo, si F es el producto punto euclidiano, el polinomio cuadrático asociado es

$$\mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^{n} x_i^2.$$

El término **homogéneo**, aplicado a un polinomio, significa que todos los términos tienen el mismo grado en las variables x_i. En toda esta sección, la expresión «polinomio cuadrático» será usada en el sentido de «polinomio cuadrático homogéneo».

La Ec. (4) puede ser escrita como el siguiente producto matricial:

$$F(\mathbf{x}, \mathbf{y}) = (x_1 \quad x_1 \quad \dots \quad x_n)\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{n1} & \dots & & a_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

o bien

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t A \mathbf{y} = \mathbf{x} \cdot A\mathbf{y}.$$

Esto se deduce inmediatamente de la definición del producto de matrices. La condición $a_{ij} = a_{ji}$ significa que la matriz A es simétrica con respecto de su diagonal.

En notación matricial, (5) tiene la forma

$$Q(\mathbf{x}) = \mathbf{x}^t A \mathbf{x},$$

y como caso especial tenemos

$$Q(\mathbf{x}) = \mathbf{x}^t I \mathbf{x} = \mathbf{x} \cdot \mathbf{x}.$$

Si $a_{ij} = 0$ para $i \neq j$, entonces A es una matriz diagonal y se dice que Q está representada en forma diagonal.

Ejemplo 1. Damos algunos ejemplos de polinomios cuadráticos.

$$(x \quad y)\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 4xy + y^2,$$

$$(x \quad y \quad z)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 + 2xz,$$

$$(x_1 \quad x_2 \quad x_3 \quad x_4)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_4^2.$$

Un polinomio cuadrático es llamado **definido positivo** si $Q(\mathbf{x}) > 0$, excepto cuando $\mathbf{x} = 0$. Nótese que si Q es definido positivo y F es la función bilineal asociada, entonces F es un producto interno en $\mathcal{R}^n$. Esto se debe a que las condiciones (1)-(3), junto con la propiedad de ser definido positivo, son las propiedades características de un producto interno.

Ejemplo 2. La Fig. 6 muestra los gráficos de dos polinomios cuadráticos. El de la izquierda es definido positivo. El otro no lo es; su gráfico se llama un **paraboloide hiperbólico.**

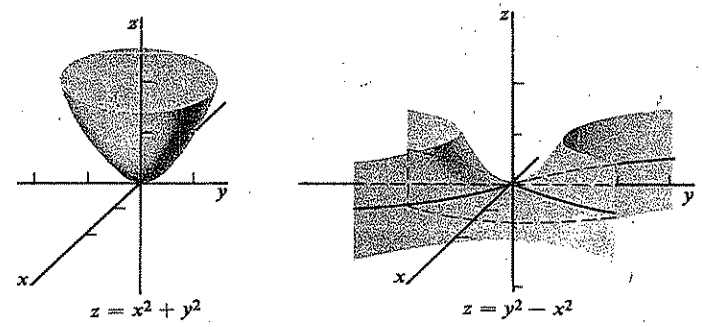


Figura 6

Ejemplo 3. El polinomio cuadrático

$$Q_1(x, y, z) = (x - y - z)^2 = x^2 + y^2 + z^2 - 2xy - 2xz + 2yz$$

es no negativo. Sin embargo, no es definido positivo, ya que él es cero en el plano $x - y - z = 0$.

El polinomio

$$Q_2(x, y, z) = (x + y + z)^2 - (x - y - z)^2 = 4xy + 4xz$$

cambia de signo. En efecto, Q_2 es negativo en el plano $x + y + z = 0$ y positivo en el plano $x - y - z = 0$, excepto en la intersección de los dos planos, donde Q_2 es cero.

El polinomio

$$Q_3(x, y, z) = x^2 + y^2$$

es no negativo, pero no es definido positivo, ya que para todo z, $Q_3(0, 0, z) = 0$.

En los ejemplos recién dados, se ilustra el hecho de que si un polinomio puede ser escrito en la forma diagonal

$$Q(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2,$$

entonces Q es definido positivo si, y solo si, todos los coeficientes a_i son positivos. Además, si algunos de los coeficientes son negativos o cero, es posible

determinar las regiones en las cuales Q es positivo o negativo. En los ejemplos que siguen, consideraremos una manera de expresar el polinomio $Q(x, y)$ en forma diagonal.

Ejemplo 4. La función bilineal simétrica más general en $\mathcal{R}^2$ se obtiene eligiendo arbitrariamente a, b y c en

$$F((x, y), (x', y')) = (x \quad y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

El polinomio cuadrático general es, por tanto,

$$Q(x, y) = ax^2 + 2bxy + cy^2.$$

Para determinar bajo qué condiciones Q es definido positivo, nótese primero que no podemos tener que $a = 0$ y $c = 0$, ya que, en ese caso, $Q(x, y) = 2bxy$ y, si $b \neq 0$, este polinomio toma valores positivo y negativo. Supongamos entonces que $a \neq 0$. Completando el cuadro, tenemos que

$$Q(x, y) = \frac{1}{a}\left[a^2\left(x + \frac{b}{a}y\right)^2 + (ac - b^2)y^2\right]. \tag{6}$$

En forma análoga, si $c \neq 0$, entonces

$$Q(x, y) = \frac{1}{c}\left[c^2\left(y + \frac{b}{c}x\right)^2 + (ac - b^2)x^2\right]. \tag{7}$$

Por tanto, Q es definido positivo si, y solo si, $ac - b^2 > 0$ y $a > 0$ o $c > 0$.

Ejemplo 5. Si hemos escrito Q en una de las formas (6) o (7), un cambio de variable evidente puede ser usado para simplificar el polinomio. Para ser específico, supongamos que $a \neq 0$ y que (6) es válido. Haciendo

$$u = x + \frac{b}{a}y, \qquad v = 0x + y,$$

podemos escribir Q en la forma

$$au^2 + \frac{1}{a}(ac - b^2)v^2.$$

Esta transformación de coordenadas corresponde a un cambio de base en el cual la base natural de $\mathcal{R}^2$ es remplazada por la base

$$\mathbf{x}_1 = (1, 0), \qquad \mathbf{x}_2 = \left(-\frac{b}{a}, 1\right).$$

Las relaciones entre las coordenadas x e y y u y v pueden ser escritas en forma matricial como

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -\frac{b}{a} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Para ver concretamente el significado geométrico de la nueva base, consideramos un ejemplo numérico. Sea

$$Q(\mathbf{x}) = Q(x, y) = x^2 + 2xy + 3y^2.$$

Entonces $a = b = 1$ y $c = 3$. La nueva base está formada por los vectores $\mathbf{x}_1 = (1, 0)$ y $\mathbf{x}_2 = (-1, 1)$. Con respecto a las nuevas coordenadas,

$$Q(\mathbf{x}) = u^2 + 2v^2,$$

donde

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Es evidente que Q es definido positivo. La Fig. 7 muestra los vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ y la curva de nivel $Q(\mathbf{x}) = 1$.

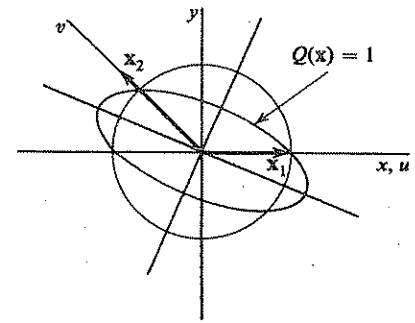


Figura 7

En el Ejemplo 5 hemos visto que un polinomio cuadrático puede ser reducido a la forma diagonal escribiéndolo en términos de coordenadas en una base apropiada. Sin embargo, esto fue hecho con vectores que no son necesariamente perpendiculares. El teorema que sigue muestra que siempre podemos encontrar una base diagonalizadora formada por vectores perpendiculares de largo 1.

2.1 Teorema

Sea Q un polinomio cuadrático en $\mathcal{R}^n$. Entonces existe una base ortonormal $\mathbf{x}_1, \ldots, \mathbf{x}_n$ tal que si $y_1, \ldots, y_n$ son las coordenadas de un vector $\mathbf{x}$ con respecto a esta base, entonces

$$Q(\mathbf{x}) = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2.$$

Por tanto,

$$Q(\mathbf{x}_k) = \lambda_k, \qquad k = 1, \ldots, n.$$

Demostración. Los vectores básicos $\mathbf{x}_1, \ldots, \mathbf{x}_n$ serán elegidos sucesivamente como sigue. Sea S^{n-1} el conjunto de todos los vectores unitarios en $\mathcal{R}^n$; es decir, el conjunto de todos los $\mathbf{x}$ tales que $|\mathbf{x}| = 1$. (El conjunto de todos los vectores unitarios de $\mathcal{R}^n$ es una superficie de dimensión $n-1$, definida implícitamente por la ecuación $|\mathbf{x}| = 1$. Por esta razón usamos el índice $n-1$ en S^{n-1}.) Sea $\mathbf{x}_1$ un punto máximo de Q restringido a S^{n-1}. (Véase la introducción al Apéndice.) Por elección, $\mathbf{x}_1$ es un vector unitario. Sea $\mathcal{U}_{n-1}$ el subespacio $(n-1)$-dimensional de $\mathcal{R}^n$ formado por los vectores $\mathbf{x}$ de $\mathcal{R}^n$ que son perpendiculares a $\mathbf{x}_1$ y sea S^{n-2} la esfera unitaria en $\mathcal{U}_{n-1}$. Restringiendo $\mathbf{x}$ a S^{n-2}, sea $\mathbf{x}_2$ un vector en S^{n-2} en el cual $Q(\mathbf{x})$ toma su valor máximo. Por elección, $\mathbf{x}_2$ es un vector unitario perpendicular a $\mathbf{x}_1$. Suponiendo que $\mathbf{x}_1, \ldots, \mathbf{x}_k$, $k < n$, han sido elegidos de esta manera, sea $\mathcal{U}_{n-k}$ el subespacio de $\mathcal{R}^n$ formado por todos los vectores perpendiculares a $\mathbf{x}_1, \ldots, \mathbf{x}_k$. Sea S^{n-k-1} la esfera unitaria en $\mathcal{U}_{n-k}$ y supongamos que Q asume su máximo en S^{n-k-1} en el vector $\mathbf{x}_{k+1}$. Continuamos este proceso hasta obtener n vectores unitarios, cada uno perpendicular a los obtenidos anteriormente.

Es claro que los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_n$ forman una base ortonormal de $\mathcal{R}^n$. Demostraremos que ésta es una base diagonalizadora para Q. Como Q tiene un máximo en $\mathbf{x}_1$ cuando se lo reduce a la esfera unitaria $|\mathbf{x}|^2 - 1 = 0$, por el teorema de Lagrange (Teorema 1.2), la función f definida por

$$f(\mathbf{x}) = Q(\mathbf{x}) - \lambda(|\mathbf{x}|^2 - 1)$$

debe tener un punto crítico en $\mathbf{x}_1$ para algún λ. Es decir, $q'(\mathbf{x}_1) = 0$. Un cálculo sencillo muestra que todo polinomio cuadrático Q en $\mathcal{R}^n$ y su función bilineal asociada F, satisfacen la ecuación

$$Q'(\mathbf{x})\mathbf{y} = 2F(\mathbf{x}, \mathbf{y})$$

para todo $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$. (Véase Ejercicio 10.) Por tanto, en el punto crítico $\mathbf{x}_1$,

$$0 = f'(\mathbf{x}_1)\mathbf{y} = 2F(\mathbf{x}_1, \mathbf{y}) - 2\lambda\mathbf{x}_1 \cdot \mathbf{y},$$

y concluimos que

$$F(\mathbf{x}_1, \mathbf{y}) = \lambda\mathbf{x}_1 \cdot \mathbf{y}$$

para todo $\mathbf{y}$ en $\mathcal{R}^n$. Por tanto,

$$F(\mathbf{x}_1, \mathbf{x}_k) = 0, \qquad k = 2, \ldots, n, \qquad F(\mathbf{x}_1, \mathbf{x}_1) = \lambda = Q(\mathbf{x}_1).$$

Restringiendo Q a $\mathcal{U}_{n-1}$, el subespacio de $\mathcal{R}^n$ perpendicular a $\mathbf{x}_1$, podemos repetir el mismo argumento y obtener que

$$F(\mathbf{x}_2, \mathbf{x}_k) = 0, \qquad k = 3, \ldots, n.$$

Continuando de esta forma, tenemos, finalmente, que

$$F(\mathbf{x}_i, \mathbf{x}_k) = 0, \qquad \text{si } i \neq k.$$

Si un vector arbitrario $\mathbf{x}$ se expresa en términos de la base $\mathbf{x}_1, \ldots, \mathbf{x}_n$ en la forma $\mathbf{x} = y_1\mathbf{x}_1 + \ldots + y_n\mathbf{x}_n$, entonces

$$Q(\mathbf{x}) = \sum_{j,k=1}^{n} F(\mathbf{x}_j, \mathbf{x}_k)y_j y_k$$

$$= \sum_{k=1}^{n} F(\mathbf{x}_k, \mathbf{x}_k)y_k^2 = \sum_{k=1}^{n} Q(\mathbf{x}_k)y_k^2.$$

Esto completa la demostración.

Otra consecuencia de la demostración es el siguiente teorema.

2.2 Teorema

Los vectores básicos $\mathbf{x}_1, \ldots, \mathbf{x}_n$ con respecto a los cuales el polinomio cuadrático Q tiene la forma

$$Q(\mathbf{x}) = \sum_{k=1}^{n} \lambda_k y_k^2$$

pueden ser elegidos con la condición que $Q(\mathbf{x}_k)$ sea el valor máximo de Q restringido a la esfera unitaria en el subespacio de $\mathcal{R}^n$ perpendicular a los vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k-1}$. La propiedad de valor máximo de los vectores básicos $\mathbf{x}_k$ puede ser usada para calcularlos, como lo hacemos en el ejemplo que sigue.

Ejemplo 6. Supongamos que el polinomio cuadrático

$$Q(x, y) = 3x^2 + 2xy + 3y^2$$

ha sido expresado en la base natural de $\mathcal{R}^2$. Restringimos Q al círculo unitario

$$x^2 + y^2 - 1 = 0.$$

Por el teorema de Lagrange (Teorema 1.2), Q tendrá su máximo en los puntos críticos de

$$3x^2 + 2xy + 3y^2 - \lambda(x^2 + y^2 - 1),$$

para algún λ. Es decir, para algún λ, el vector (x, y) debe satisfacer

$$(3 - \lambda)x + y = 0$$
$$x + (3 - \lambda)y = 0$$

además de $x^2 + y^2 = 1$. (λ ha sido remplazado por $-\lambda$.) Estas ecuaciones tendrán una solución distinta de cero solo si las columnas de la matriz

$$\begin{pmatrix} 3 - \lambda & 1 \\ 1 & 3 - \lambda \end{pmatrix}$$

son dependientes. Como la dependencia es equivalente a

$$\begin{vmatrix} 3 - \lambda & 1 \\ 1 & 3 - \lambda \end{vmatrix} = 0,$$

debemos tener que $(3 - \lambda)^2 - 1 = 0$, es decir, $\lambda = 2, 4$. Las soluciones correspondientes para (x, y) son

$$\lambda = 2: \quad (x, y) = \left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}\right).$$

$$\lambda = 4: \quad (x, y) = \left(\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right).$$

El máximo de Q se obtiene en $(\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$, de modo que podemos elegir $\mathbf{x}_1 = (1/\sqrt{2}, 1/\sqrt{2})$. Para $\mathbf{x}_2$ podemos tomar uno de los dos vectores $(-1/\sqrt{2}, 1/\sqrt{2})$ o su negativo.

Sea $\mathbf{x}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. La ecuación de cambio de coordenadas es, por tanto,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

En términos de las nuevas variables, tenemos

$$Q(\mathbf{x}) = 4u^2 + 2v^2.$$

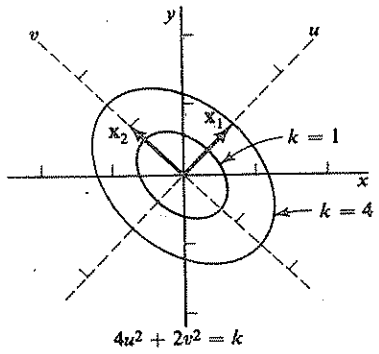


Figure 8

La Fig. 8 muestra las curvas de nivel de Q y su relación con los vectores básicos originales y con los nuevos.

Para algunos fines, no es necesario encontrar la base ortonormal $\mathbf{x}$ del Teorema 2.1, siempre que se puedan calcular los números λ. Por ejemplo, es claro

que si conocemos λ_k podemos decidir si Q es definido positivo o no. El teorema siguiente nos permite calcular o estimar los λ_k.

2.3 Teorema

Sea Q un polinomio cuadrático en $\mathcal{R}^n$ dado por

$$Q(\mathbf{x}) = \mathbf{x}^t A \mathbf{x},$$

donde A es una matriz simétrica. Supongamos que, con respecto de las coordenadas en la base ortonormal $\mathbf{x}_1, \ldots, \mathbf{x}_n$, Q tiene la forma

$$Q(\mathbf{x}) = \sum_{k=1}^{n} \lambda_k y_k^2, \tag{8}$$

donde $Q(\mathbf{x}_k) = \lambda_k$. Entonces los números λ_k son las raíces de la ecuación

$$\det (A - \lambda I) = 0. \tag{9}$$

A pesar de que la existencia de los vectores básicos $\mathbf{x}_1, \ldots, \mathbf{x}_n$ ha sido demostrada en el Teorema 2.1, no es necesario conocerlos para determinar los λ_k. Ellos pueden ser calculados resolviendo la Ec. (9), llamada la **ecuación característica** de Q. Las raíces λ_k de esta ecuación se llaman las **raíces características** o **valores propios**. El teorema que sigue nos da otro método para calcular los vectores básicos $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

2.4 Teorema

Sea $\mathbf{z}_1, \ldots, \mathbf{z}_n$ una base ortonormal tal que para todo $k = 1, \ldots, n$, el vector $\mathbf{z}_k$ satisface la ecuación matricial

$$(A - \lambda_k I)\mathbf{z}_k = 0. \tag{10}$$

Entonces Q tiene la forma diagonal (8).

Vectores $\mathbf{z}_k$ que satisfacen la Ec. (10) se llaman **vectores característicos** o **propios correspondientes** a λ_k.

Demostración (del Teorema 2.3). Supongamos que los vectores básicos ortonormales que diagonalizan Q son

$$\mathbf{x}_1 = \begin{pmatrix} b_{11} \\ \cdot \\ \cdot \\ \cdot \\ b_{n1} \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} b_{12} \\ \cdot \\ \cdot \\ \cdot \\ b_{n2} \end{pmatrix}, \quad \ldots, \quad \mathbf{x}_n = \begin{pmatrix} b_{1n} \\ \cdot \\ \cdot \\ \cdot \\ b_{nn} \end{pmatrix}.$$

Sea B la matriz n-por-n con columnas $\mathbf{x}_1, \ldots, \mathbf{x}_n$. De acuerdo con la Sec. 10 del Cap. 2, las coordenadas de un mismo punto están relacionadas por la ecuación $\mathbf{x} = B\mathbf{y}$, donde

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad \text{y} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix},$$

$y_1, \ldots, y_n$ son las coordenadas con respecto a $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Entonces sustituyendo $\mathbf{x}$ por $B\mathbf{y}$, tenemos

$$\begin{aligned} Q(\mathbf{x}) &= (B\mathbf{y})^t A(B\mathbf{y}) \\ &= \mathbf{y}^t(B^t AB)\mathbf{y} \\ &= \mathbf{y}^t \Lambda \mathbf{y}. \end{aligned}$$

Por elección de las columnas de B, la matriz $\Lambda = B^t AB$ es diagonal con elementos $\lambda_1, \ldots, \lambda_n$ en su diagonal. Además, como las columnas de B son las coordenadas de vectores unitarios perpendiculares con respecto a una base ortonormal, tenemos que $B^t B = I$ como se comprueba por multiplicación. En otras palabras, $B^t = B^{-1}$. Entonces $\Lambda = B^{-1}AB$. Si restamos λI a ambos lados de la ecuación y factorizamos el miembro de la derecha, tenemos

$$\begin{aligned} \Lambda - \lambda I &= B^{-1}AB - \lambda I \\ &= B^{-1}(A - \lambda I)B. \end{aligned}$$

Pero $\Lambda - \lambda I$ es una matriz diagonal con elementos $\lambda_k - \lambda$ en la diagonal, de modo que

$$\begin{aligned} (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda) &= \det(\Lambda - \lambda I) \\ &= \det B^{-1} \det(A - \lambda I) \det B \\ &= \det(A - \lambda I). \end{aligned}$$

Esto muestra que las raíces de $\det(A - \lambda I) = 0$ son $\lambda_1, \ldots, \lambda_n$.

Demostración (del Teorema 2.4). Sea $\mathbf{z}_1, \ldots, \mathbf{z}_n$ una base ortonormal tal que $A\mathbf{z}_k = \lambda_k \mathbf{z}_k$ para $k = 1, 2, \ldots, n$. Sea C la matriz con columnas $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$. La ecuación

$$\mathbf{x} = C\mathbf{z}$$

expresa la relación entre las coordenadas de la base $\mathbf{z}_1, \ldots, \mathbf{z}_n$ en $\mathcal{R}^n$ y las coordenadas naturales. Entonces

$$Q(\mathbf{x}) = \mathbf{x}^t A\mathbf{x} = (C\mathbf{z})^t A(C\mathbf{z}) = \mathbf{z}^t(C^t AC)\mathbf{z}.$$

Lo único que tenemos que hacer es verificar que la matriz $C^t AC$ es diagonal con elementos diagonales $\lambda_1, \ldots, \lambda_n$. Esquemáticamente, escribimos

$$\begin{aligned} C^t AC &= \begin{pmatrix} \mathbf{z}_1^t \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{z}_n^t \end{pmatrix} A(\mathbf{z}_1 \quad \ldots \quad \mathbf{z}_n) \\ &= \begin{pmatrix} \mathbf{z}_1^t \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{z}_n^t \end{pmatrix} (A\mathbf{z}_1 \quad \ldots \quad A\mathbf{z}_n) \\ &= \begin{pmatrix} \mathbf{z}_1^t \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{z}_n^t \end{pmatrix} (\lambda_1 \mathbf{z}_1 \quad \ldots \quad \lambda_n \mathbf{z}_n). \end{aligned}$$

Usando las relaciones

$$\mathbf{z}_i^t \mathbf{z}_j = \mathbf{z}_i \cdot \mathbf{z}_j = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j, \end{cases}$$

obtenemos que

$$C^t AC = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & & & \lambda_n \end{pmatrix}.$$

Esto termina la demostración.

Ejemplo 7. Sea $Q(x, y, z) = xy + yz + zx$. La matriz de Q es

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

y la ecuación característica es

$$\begin{pmatrix} -\lambda & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\lambda \end{pmatrix} = 0$$

es decir,

$$-\lambda^3 + \tfrac{3}{4}\lambda + \tfrac{1}{4} = 0.$$

Las raíces características son $\lambda = 1, -\frac{1}{2}, -\frac{1}{2}$, de modo que existe un sistema ortonormal de coordenadas (u, v, w) con respecto a las cuales Q tiene la forma

$$u^2 - \tfrac{1}{2}v^2 - \tfrac{1}{2}w^2.$$

Para encontrar los vectores básicos asociados, buscamos vectores unitarios que son solución de las ecuaciones

$$\begin{pmatrix} -\lambda & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

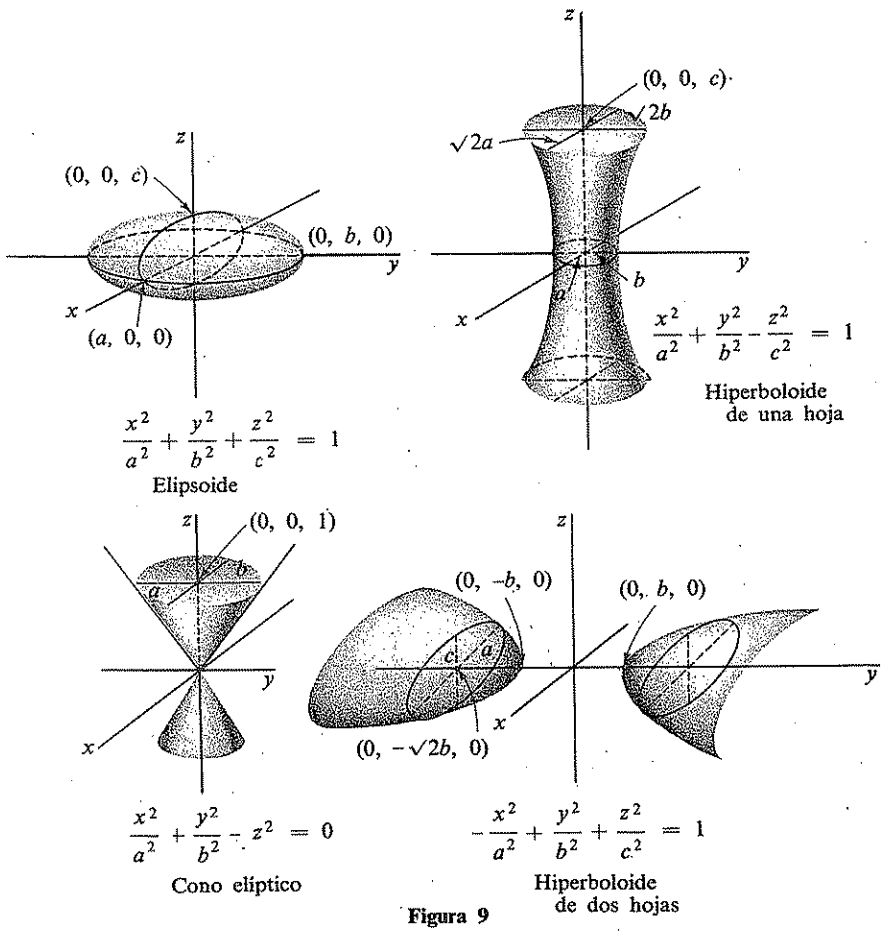


Figura 9

con $\lambda = 1$ y $\lambda = -\frac{1}{2}$. Con $\lambda = 1$, obtenemos la solución $x = y = z$, de modo que podemos elegir $x_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. Cuando $\lambda = \frac{1}{2}$, la ecuación matricial exige únicamente que los dos vectores básicos que restan estén en el plano perpendicular a x_1, de ecuación $x + y + z = 0$. Entonces, $\mathbf{x}_2$ y $\mathbf{x}_3$ pueden ser dos vectores perpendiculares cualesquiera en este plano, por ejemplo,

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right) \quad \text{y} \quad \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right).$$

Una superficie de nivel de un polinomio cuadrático se llama **superficie cuadrática**. Como todo polinomio cuadrático puede ser diagonalizado con respecto de alguna base ortonormal, es suficiente conocer las superficies de nivel de algunos polinomios cuadráticos estándar para poder conocer las superficies cuadráticas en general. La Fig. 9 muestra algunas superficies cuadráticas en $\mathcal{R}^3$.

Ejemplo 8. El polinomio cuadrático $xy + yz$ puede ser representado por una matriz simétrica en la forma

$$xy + yz = (x \quad y \quad z)\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

La ecuación característica es $-\lambda^3 + \frac{1}{2}\lambda = 0$, que tiene las raíces $\lambda = 1/\sqrt{2}$, 0, $-1/\sqrt{2}$. Las ecuaciones correspondientes para los vectores característicos y sus vectores unitarios solución son las siguientes:

$$\lambda = \frac{1}{\sqrt{2}}: \quad \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

$$(x, y, z) = \pm\left(\frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{2}\right)$$

$$\lambda = 0: \quad \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0, \qquad (x, y, z) = \pm\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

$$\lambda = -\frac{1}{\sqrt{2}}: \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

$$(x, y, z) = \pm\left(\frac{1}{2}, -\frac{1}{\sqrt{2}}, \frac{1}{2}\right).$$

Escogemos la base diagonalizadora

$$\mathbf{x}_1 = \left(\frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{2}\right), \quad \mathbf{x}_2 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \quad \mathbf{x}_3 = \left(\frac{1}{2}, -\frac{1}{\sqrt{2}}, \frac{1}{2}\right).$$

Si u, v y w son las coordenadas con respecto a la nueva base, entonces el polinomio tendrá la forma $(1/\sqrt{2})u^2 - (1/\sqrt{2})w^2$. La superficie de nivel

$$\frac{1}{\sqrt{2}}u^2 - \frac{1}{\sqrt{2}}w^2 = 1$$

es el cilindro hiperbólico, una hoja del cual es mostrada en la Fig. 10.

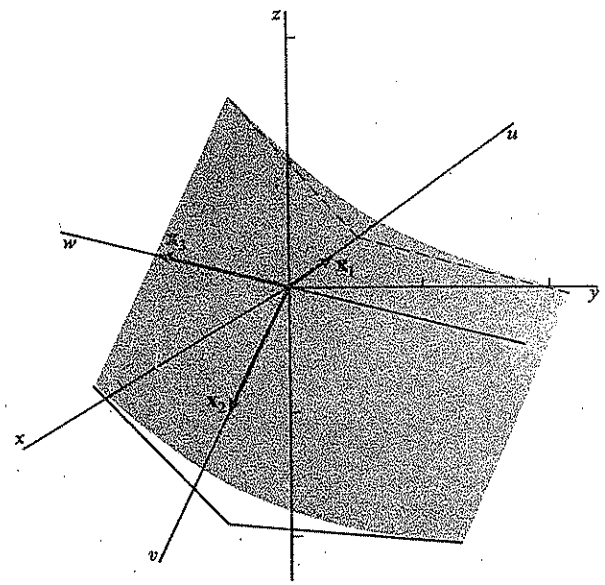


Figura 10

Los polinomios homogéneos de grado $N > 2$ pueden ser definidos en forma análoga a los polinomios cuadráticos. Podemos empezar por una función real $F(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ definida para cualquier N-pla de vectores $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$, simétrica en sus N variables vectoriales y lineal en cada una de ellas. Entonces

$$P(\mathbf{x}) = F(\mathbf{x}, \mathbf{x}, \ldots, \mathbf{x})$$

es un polinomio de grado N. Otra manera de hacer esto es considerar funciones de $\mathbf{x} = (x_1, \ldots, x_n)$ que tengan la forma

$$P(\mathbf{x}) = \sum_{i_1, \ldots, i_N = 1}^{n} a_{i_1 \ldots i_N} x_{i_1} \cdots x_{i_N},$$

donde los coeficientes $a_{i_1, \ldots, i_N}$ son simétricos en los subíndices. En ambos casos obtenemos las mismas funciones. La condición de simetría de F o de los coeficientes es introducida por conveniencia y, si no se introdujera, se podría obtener simplemente tomando el valor medio de los coeficientes no simétricos. Los detalles son dejados como ejercicio (véase Ejercicio 12).

EJERCICIOS

1. Exprese los polinomios cuadráticos como una suma de cuadrados, usando un cambio de coordenadas. En cada problema encuentre una base ortonormal que realice este cambio de coordenadas y encuentre las transformaciones de coordenadas.

(a) $3x^2 + 2\sqrt{2}\,xy + 4y^2$.

$$\left[\textit{Resp}: Q(\mathbf{x}) = 2u^2 + 5v^2; \mathbf{x}_1 = \left(\frac{\sqrt{2}}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right), \mathbf{x}_2 = \left(\frac{1}{\sqrt{3}}, \frac{\sqrt{2}}{\sqrt{3}}\right).\right]$$

(b) $3x^2 + 2\sqrt{3}\,xy + 5y^2$.

$$\left[\textit{Resp}: Q(\mathbf{x}) = 2u^2 + 6v^2; \mathbf{x}_1 = \left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \mathbf{x}_2 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right).\right]$$

(c) $(x \quad y)\begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$.

(d) $2x^2 - 5xy + 2y^2 - 2xz + 4z^2 - 2yz$.

$$\left[\textit{Resp}: Q(\mathbf{x}) = \frac{9}{2}u^2 - \frac{9}{2}v^2 + 0w^2; \quad \mathbf{x}_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \right.$$
$$\left. \mathbf{x}_2 = \left(\frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{4}{3\sqrt{2}}\right), \quad \mathbf{x}_3 = \left(\frac{3}{2}, \frac{2}{3}, -\frac{1}{3}\right).\right]$$

(e) $(x \quad y \quad z)\begin{pmatrix} -1 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

2. (a) Para cada polinomio Q del Ejercicio 1, encuentre el máximo de Q restringido a la esfera unitaria $|\mathbf{x}| = 1$ del espacio euclidiano en que Q está definida. [*Indicación*: Véase el Teorema 2.2.]

(b) Encuentre el máximo del polinomio en 1(a) restringido al círculo $x^2 + y^2 = 3$. [*Resp.* 15.]

3. Clasifique los polinomios cuadráticos siguientes como positivos definidos, negativos definidos o ninguno. Dé razones. (Q es **definida negativa** si $Q < 0$ excepto por $Q(0) = 0$.)

(a) $2x^2 - 7xy + 5y^2$. [*Resp.* Ninguna.]

(b) $2x^2 - 3xy + 5y^2$. [*Resp.* Definido positivo.]

(c) $-x^2 + 2xy - 6y^2$. [*Resp.* Definido negativo.]

(d) $3x^2 + xy + 3y^2 + 5z^2$. [*Resp.* Definido positivo.]

(e) $(x \quad y \quad z)\begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. [*Resp.* Ninguno.]

4. Demuestre que $x^2 + y^2 + z^2 - xy - xz - yz$ no es definida positiva, pero que sí lo es cuando se la restringe al plano $x + y + z = 0$.

5. Dibuje las curvas de nivel $Q(\mathbf{x}) = 1$ y $Q(\mathbf{x}) = 0$ para cada uno de los siguientes polinomios en $\mathcal{R}^2$.

(a) xy. (b) $x^2 + xy + y^2$. (c) $x^2 + xy - 2y^2$.

6. Dibuje las superficies de nivel $Q(\mathbf{x}) = 1$ y $Q(\mathbf{x}) = 0$ para los siguientes polinomios en $\mathcal{R}^3$.

(a) $x^2 - xy + y^2 + z^2$. (b) $x^2 + xy$. (c) $x^2 - 2xy + y^2 - z^2$.

7. Demuestre que todo polinomio cuadrático Q satisface

$$Q(a\mathbf{x}) = a^2 Q(\mathbf{x}),$$

para todo número real a.

8. Sea Q un polinomio cuadrático en $\mathcal{R}^n$ y F su función bilineal simétrica asociada. Demuestre que

$$F(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})],$$

para todos los vectores $\mathbf{x}$ e $\mathbf{y}$ en $\mathcal{R}^n$. Esta ecuación demuestra que F está determinada únicamente por Q.

9. Demuestre que todo polinomio cuadrático en $\mathcal{R}^n$ es una función continua.

10. Sea Q un polinomio cuadrático en $\mathcal{R}^n$ y F su función bilineal simétrica asociada, es decir, $Q(\mathbf{x}) = F(\mathbf{x}, \mathbf{x})$. Demuestre que Q es una función continuamente diferenciable o, más explícitamente, que

$$Q'(\mathbf{x})\mathbf{y} = 2F(\mathbf{x}, \mathbf{y}).$$

11. Demuestre que todo polinomio cuadrático es una función continuamente diferenciable.

12. Lo que sigue ilustra el hecho de que la condición de simetría para una función bilineal puede ser obtenida eliminando la falta de simetría tomando el promedio. Sea G una función real definida para todo par de vectores $\mathbf{x}$, $\mathbf{y}$ que es lineal en cada variable (*no* suponemos que G es simétrica). Demuestre que la función F definida por

$$F(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}(G(\mathbf{x}, \mathbf{y}) + G(\mathbf{y}, \mathbf{x}))$$

es una función bilineal simétrica. Demuestre que $G(\mathbf{x}, \mathbf{x}) = F(\mathbf{x}, \mathbf{x})$ de modo que G y F definen el mismo polinomio cuadrático.

13. Sea Q un polinomio cuadrático en $\mathcal{R}^n$. Demuestre que existe una base $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de $\mathcal{R}^n$ tal que, para cualquier vector $\mathbf{x} = y_1\mathbf{x}_1 + \ldots + y_n\mathbf{x}_n$,

$$Q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i y_i^2, \qquad \text{con} \quad \lambda_i = 0, 1, \text{ o } -1.$$

14. Demuestre que si Q es un polinomio cuadrático positivo definido en $\mathcal{R}^n$, entonces existe un número real positivo m tal que

$$Q(\mathbf{x}) \geq m|\mathbf{x}|^2, \qquad \text{para todo } \mathbf{x} \text{ en } \mathcal{R}^n.$$

[*Sugerencia*: Diagonalice Q.] Un corolario es que los valores de Q en la esfera unitaria $|\mathbf{x}| = 1$ son mayores que cero.

15. Sea Q un polinomio cuadrático positivo definido en $\mathcal{R}^2$ y sean λ_1 y λ_2 sus raíces características. Demuestre que $\lambda_1^{-1/2}$ y $\lambda_2^{-1/2}$ son las longitudes de los ejes principales de la elipse $Q(\mathbf{x}) = 1$.

16. Verifique que si $a \neq 0$ y $ab - f^2 \neq 0$, entonces

$$(x \quad y \quad z)\begin{pmatrix} a & f & e \\ f & b & d \\ e & d & c \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = a\left(x + \frac{f}{a}y + \frac{e}{a}z\right)^2 + \frac{\begin{vmatrix} a & f \\ f & b \end{vmatrix}}{a}\left(y + \frac{\begin{vmatrix} a & f \\ e & d \end{vmatrix}}{\begin{vmatrix} a & f \\ f & b \end{vmatrix}}z\right)^2 + \frac{\begin{vmatrix} a & f & e \\ f & b & d \\ e & d & c \end{vmatrix}}{\begin{vmatrix} a & f \\ f & b \end{vmatrix}}z^2.$$

Deduzca de esto que este polinomio es definido positivo si, y solo si, los tres determinantes siguientes son positivos:

$$\begin{vmatrix} a & f & e \\ f & b & d \\ e & d & c \end{vmatrix}$$

¿Cuál es la condición para que el polinomio sea negativo definido? Este criterio puede ser extendido a polinomios cuadráticos de $\mathcal{R}^n$. Véase R. M. Thrall y L. Tornheim, *Vector Spaces and Matrices*, Wiley, 1957, pág. 170.

17. Demuestre que la transformación lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^n$ determinada por una matriz simétrica A tiene la propiedad que

$$L(\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot L(\mathbf{y})$$

para todo x e y en $\mathcal{R}^n$. Conversamente, demuestre que si esta ecuación es satisfecha, entonces A es una matriz simétrica.

18. Una transformación $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^n$ para la cual es válida la ecuación del Ejercicio 17, se llama una transformación simétrica. Demuestre que para toda transformación simétrica existe una base ortonormal tal que, con respecto a esta base, la matriz de la transformación es diagonal.

19. (a) Dé una interpretación geométrica de la acción en $\mathcal{R}^3$ de las transformaciones simétricas con matrices

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

(b) Use el resultado del Ejercicio 18 para dar una interpretación geométrica de la acción de una transformación simétrica sobre $\mathcal{R}^3$.

20. La demostración del Teorema 2.1 funciona también cuando se remplaza el producto interno euclidiano por uno cualquiera. Use este hecho para demostrar el siguiente teorema. Si Q_1 y Q_2 son polinomios cuadráticos en $\mathcal{R}^n$ y Q_1 es positivo definido, entonces existe una base ortonormal para $\mathcal{R}^n$ con respecto a la cual Q_1 y Q_2 tienen forma diagonal.

21. Demuestre el converso del Teorema 2.4, es decir, que si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ es una base ortonormal que diagonaliza a $Q(\mathbf{x}) = \mathbf{x}^t A\mathbf{x}$, entonces $(A - \lambda_k I)\mathbf{x}_k = 0$, donde $\lambda_k = Q(\mathbf{x}_k)$.

3. Desarrollos de Taylor

Empezaremos por recordar la definición y las propiedades más sencillas del desarrollo de Taylor para funciones de una variable. Si $f(x)$ tiene una derivada N-ésima en x_0, su **desarrollo de Taylor de grado** N alrededor de x_0 es el polinomio

$$f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \ldots + \frac{1}{N!}f^{(N)}(x_0)(x - x_0)^N.$$

La relación entre f y su desarrollo de Taylor puede ser expresada convenientemente mediante la **fórmula integral del resto.**

3.1 Teorema

Si f tiene una derivada N-ésima continua en una vecindad de x_0, entonces, en esa vecindad,

$$f(x) = f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \ldots + \frac{1}{N!}f^{(N)}(x_0)(x - x_0)^N + R_N, \quad (1)$$

donde

$$R_N = \frac{1}{(N-1)!}\int_{x_0}^{x}(x - t)^{N-1}[f^{(N)}(t) - f^{(N)}(x_0)]\,dt.$$

Demostración. El resto puede ser expresado como la diferencia

$$R_N = \frac{1}{(N-1)!}\int_{x_0}^{x}(x - t)^{N-1}f^{(N)}(t)\,dt - \frac{f^{(N)}(x_0)}{(N-1)!}\int_{x_0}^{x}(x - t)^{N-1}dt.$$

La segunda integral puede ser calculada directamente, obteniéndose

$$\frac{1}{N!}f^{(N)}(x_0)(x - x_0)^N,$$

el cual es exactamente el último término del desarrollo de Taylor. Integrando por partes la primera integral, se tiene

$$\frac{1}{(N-2)!}\int_{x_0}^{x}(x - t)^{N-2}[f^{(N-1)}(t) - f^{(N-1)}(x_0)]\,dt = R_{N-1}.$$

Por tanto, obtenemos

$$R_N = -\frac{1}{N!}f^{(N)}(x_0)(x - x_0)^N + R_{N-1}.$$

Si sustituimos la ecuación precedente en (1), obtenemos (1) de nuevo con N remplazado por $N - 1$. La inducción se completa notando que para $N = 1$, la Ec. (1) es simplemente

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x}[f'(t) - f'(x_0)]\,dt,$$

la cual es verdadera.

De la fórmula del resto se deduce que un polinomio de grado N es igual a su desarrollo de Taylor de grado N. En efecto, si f es de grado N, entonces $f^{(N)}$ es una función constante, de modo que el resto es idénticamente cero. A continuación damos una lista de ejemplos comunes. Solo en el primero tenemos una igualdad. Los desarrollos son todos alrededor de $x = 0$.

$$(1 + x)^N = \sum_{k=0}^{N}\binom{N}{k}x^k,$$

$$\frac{1}{(1-x)^M}: \quad \sum_{k=0}^{N}\binom{M + k - 1}{M - 1}x^k,$$

$$e^x: \quad 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \ldots + \frac{1}{N!}x^N,$$

$$\log(1 - x): \quad -x - \tfrac{1}{2}x^2 - \tfrac{1}{3}x^3 - \ldots - \frac{1}{N}x^N, \quad (2)$$

$$\cos x: \quad 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^k \frac{x^{2k}}{(2k)!},$$

$$\text{sen } x: \quad x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

La Fig. 11 muestra gráficamente la relación entre e^x y cos x y sus desarrollos de Taylor de grado 2.

Para ver cómo generalizar el desarrollo de Taylor a las funciones de más de una variable, supongamos primero que tenemos una función $f(x, y)$ con derivadas parciales en (x_0, y_0) hasta el orden N. Primero consideremos la función F de una variable definida por

$$F(t) = f(tu + x_0, tv + y_0),$$

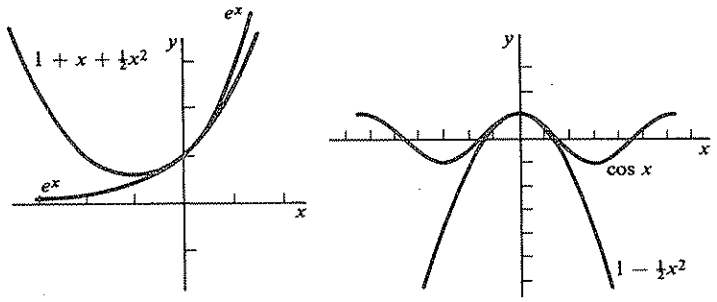


Figura 11

donde u y v mantienen fijos. Trataremos de encontrar un desarrollo de Taylor de F en $t = 0$. Para ello calculamos las derivadas de F, obteniendo, por la regla de la cadena, que

$$F'(t) = u \frac{\partial f}{\partial x}(tu + x_0, ty + y_0) + v \frac{\partial f}{\partial y}(tv + x_0, ty + y_0),$$

y, por tanto, que

$$F'(0) = u \frac{\partial f}{\partial x}(x_0, y_0) + v \frac{\partial f}{\partial y}(x_0, y_0).$$

En forma análoga, calculamos $F''(t)$ y, haciendo $t = 0$, obtenemos

$$F''(0) = u^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2uv \frac{\partial^2 f}{\partial x\, \partial y}(x_0, y_0) + v^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

Siguiendo de esta manera, obtenemos

$$F^{(N)}(0) = \sum_{k=0}^{N} \binom{N}{k} u^k v^{N-k} \frac{\partial^k f}{\partial x^k}(x_0, y_0) \frac{\partial^{N-k} f}{\partial y^{N-k}}(x_0, y_0),$$

donde $\binom{N}{k}$ es el coeficiente binomial $N!/k!(N-k)!$ Entonces, si remplazamos u por $(x - x_0)$ y v por $(y - y_0)$, tenemos el desarrollo de Taylor

$$F(t) \sim F(0) + \frac{1}{1!} F'(0)t + \frac{1}{2!} F''(0)t^2 + \ldots$$

que, para $t = 1$, da

$$\begin{aligned} f(x, y) \sim f(x_0, y_0) &+ \frac{1}{1!}\left(\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)\right) \\ &+ \frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x - x_0)^2 + 2\frac{\partial^2 f}{\partial x\, \partial y}(x - x_0)(y - y_0)\right. \\ &\left. + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y - y_0)^2\right) + \ldots. \end{aligned}$$

Esto nos lleva a hacer la siguiente definición. Para una función de dos variables $f(x, y)$ con derivadas parciales de orden N continuas en una vecindad de (x_0, y_0), el **desarrollo de Taylor de grado N alrededor de** (x_0, y_0) es el polinomio

$$\begin{aligned} f(x_0, y_0) &+ \frac{1}{1!}\left((x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0)\right) \\ &+ \frac{1}{2!}\left((x - x_0)^2 f_{xx}(x_0, y_0) + 2(x - x_0)(y - y_0) f_{xy}(x_0, y_0)\right. \\ &\qquad \left. + (y - y_0)^2 f_{yy}(x_0, y_0)\right) \quad (3) \\ &\vdots \\ &+ \frac{1}{N!} \sum_{k=0}^{N} \binom{N}{k} (x - x_0)^k (y - y_0)^{N-k} f_{x^k y^{N-k}}(x_0, y_0). \end{aligned}$$

Ejemplo 1. Sea $f(x, y) = \sqrt{1 + x^2 + y^2}$. Para obtener el desarrollo de grado 2 alrededor de (0, 0) calcularemos

$$f_x(0, 0) = f_y(0, 0) = 0,$$
$$f_{xx}(0, 0) = f_{yy}(0, 0) = 1 \quad \text{y} \quad f_{xy}(0, 0) = 0.$$

Entonces la Ec. (3) se reduce al polinomio de segundo grado $1 + \frac{1}{2}(x^2 + y^2)$. Los gráficos de f y de su desarrollo de Taylor de segundo grado son mostrados en la Fig. 12.

Para simplificar la escritura de los términos del desarrollo de Taylor, podemos usar la notación siguiente. El operador diferencial

$$\left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}\right)$$

aplicado a f y evaluado en $\mathbf{x}_0 = (x_0, y_0)$, es, por definición, el polinomio

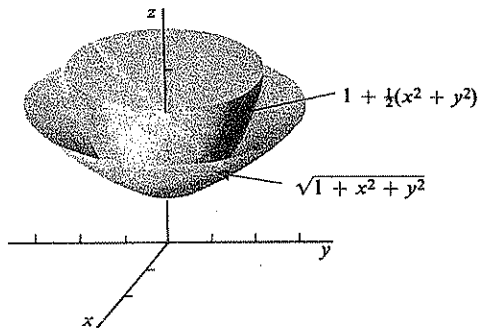


Figura 12

$$d_{\mathbf{x}_0} f(x, y) = \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}\right)_{\mathbf{x}_0} f = x \frac{\partial f}{\partial x}(\mathbf{x}_0) + y \frac{\partial f}{\partial y}(\mathbf{x}_0).$$

También se pueden definir diferenciales de orden k, $k > 1$. Ellas son polinomios homogéneos de grado k. Si f tiene las derivadas necesarias, entonces su definición es

$$d^k_{\mathbf{x}_0} f(x, y) = \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}\right)^k_{\mathbf{x}_0} f = \sum_{j=0}^{k} \binom{k}{j} x^j y^{k-j} \frac{\partial^k f}{\partial x^j \partial y^{k-j}}(\mathbf{x}_0).$$

En estas fórmulas, el operador

$$\left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}\right)^k$$

ha sido multiplicado de acuerdo con el desarrollo del binomio. El operador se aplica a f y entonces las derivadas parciales se evalúan en $\mathbf{x}_0 = (x_0, y_0)$. Nótese que x e y son las únicas variables que aparecen en la ecuación anterior, ya que

$$\binom{k}{j} \frac{\partial^k f}{\partial x^j \partial y^{k-j}}(\mathbf{x}_0)$$

es una constante para $\mathbf{x}_0$ fijo. El desarrollo de Taylor de grado N de f en $\mathbf{x}_0$ puede ser escrito ahora en la forma

$$f(x_0, y_0) + \frac{1}{1!} d_{\mathbf{x}_0} f(x - x_0, y - y_0) + \ldots + \frac{1}{N!} d^N_{\mathbf{x}_0} f(x - x_0, y - y_0).$$

Ejemplo 2. Si $f(x, y)$ tiene las derivadas necesarias, entonces la diferencial de orden 3 en (1, 1), evaluada en $(x - 1, y - 1)$, es el polinomio

$$\begin{aligned} d^3_{(1,1)} f(x - 1, y - 1) &= \left((x - 1)\frac{\partial}{\partial x} + (y - 1)\frac{\partial}{\partial y}\right)^3_{(1,1)} f \\ &= (x - 1)^3 \frac{\partial^3 f}{\partial x^3}(1, 1) \\ &\quad + 3(x - 1)^2 (y - 1) \frac{\partial^3 f}{\partial x^2 \partial y}(1, 1) \\ &\quad + 3(x - 1)(y - 1)^2 \frac{\partial^3 f}{\partial x \partial y^2}(1, 1) \\ &\quad + (y - 1)^3 \frac{\partial^3 f}{\partial y^3}(1, 1). \end{aligned}$$

Ejemplo 3. Si se efectúan las multiplicaciones en el polinomio $(x_1 + x_2 + \ldots + x_n)^N$, cada término estará formado por una constante multiplicando a un factor de la forma $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, donde los términos no negativos k_i satisfacen $k_1 + \ldots + k_n = N$. El desarrollo multinomial tiene la forma

$$(x_1 + \ldots + x_n)^N = \sum_{k + \ldots + k_n = N} \binom{N}{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n}.$$

Los coeficientes multinomiales pueden ser calculados y se tiene que

$$\binom{N}{k_1 \ldots k_n} = \frac{N!}{k_1! \ldots k_n!}.$$

Este cálculo se hará más adelante usando el teorema de Taylor (Teorema 3.2). Los coeficientes también pueden ser calculados contando (véase Kemeny, Snell, Thompson, *Finite Mathematics*, Prentice-Hall, 1966, pág. 109).

Para una función de n variables, la **diferencial de orden** k **en** $\mathbf{x}_0 = (a_1, \ldots, a_n)$ es, por definición, el siguiente polinomio en $\mathbf{x} = (x_1, \ldots, x_n)$:

$$\begin{aligned} d^k_{\mathbf{x}_0} f(\mathbf{x}) &= \left(x_1 \frac{\partial}{\partial x_1} + \ldots + x_n \frac{\partial}{\partial x_n}\right)^k_{\mathbf{x}_0} f \\ &= \sum_{k_1 + \ldots + k_n = k} \binom{k}{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n} \frac{\partial^k f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}(a_1, \ldots, a_n). \end{aligned}$$

En términos de las diferenciales, el desarrollo de Taylor alrededor de $\mathbf{x}_0$ es, por definición,

$$f(\mathbf{x}_0) + \frac{1}{1!} d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2!} d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + \ldots - \frac{1}{N!} d^N_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0).$$

La función $d_{\mathbf{x}_0} f$ es la misma diferencial definida en la Sec. 8 del Cap. 3. Para completar, podemos definir la diferencial de orden 0 por

$$d^0_{\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Ejemplo 4. El desarrollo de Taylor de segundo grado de $e^{x_1+\ldots+x_n}$ alrededor de $\mathbf{x} = 0$ es

$$1 + \frac{1}{1!}\left(x_1\frac{\partial}{\partial x_1} + \ldots + x_n\frac{\partial}{\partial x_n}\right)_0 f + \frac{1}{2!}\left(x_1\frac{\partial}{\partial x_1} + \ldots + x_n\frac{\partial}{\partial x_n}\right)_0^2 f$$
$$= 1 + \frac{1}{1!}(x_1 + \ldots + x_n) + \frac{1}{2!}(x_1^2 + 2x_1x_2 + \ldots + x_2^2 + \ldots + x_n^2).$$

De acuerdo con los párrafos anteriores, el desarrollo de Taylor de la función f está definido de modo que los coeficientes del polinomio puedan ser calculados en forma sencilla a partir de las derivadas de f. El desarrollo de Taylor es importante, ya que nos da una aproximación polinomial de f cerca de $\mathbf{x}_0$ que muestra de una manera sencilla muchas de las características de f cerca de $\mathbf{x}_0$. Además, a medida que se incluyen más términos de grado superior en este desarrollo, la aproximación se hace mejor. Consideremos primero el caso de una variable. El desarrollo

$$f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) = f(x_0) + d_{x_0}f(x - x_0)$$

es la aproximación afín de f dada por $d_{x_0}f$. En otras palabras, si x tiende a x_0 la función

$$f(x) - f(x_0) - f'(x_0)(x - x_0)$$

tiende a cero más rápidamente que $x - x_0$.

Una vez encontrada la aproximación de primer grado, buscamos una de segundo grado. El próximo teorema afirma que la aproximación deseada es el desarrollo de Taylor de segundo grado y que si x tiende a x_0,

$$f(x) - f(x_0) - \frac{1}{1!}f'(x_0)(x - x_0) - \frac{1}{2!}f''(x_0)(x - x_0)^2$$

tiende a cero más rápidamente que $(x - x_0)^2$. Para una función de dos variables, el desarrollo de Taylor de primer grado es

$$f(x_0, y_0) + \frac{1}{1!}\left(\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)\right)$$
$$= f(x_0, y_0) + (d_{\binom{x_0}{y_0}}f)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

y es, por tanto, la aproximación afín de f definida por la diferencial. Veremos que para encontrar una aproximación similar de segundo grado, deberemos considerar el desarrollo de Taylor de segundo grado. Damos a continuación el resultado completo, cuya demostración será dada en el Apéndice.

3.2 Teorema de Taylor

Sea $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ una función con derivadas continuas hasta de orden N en una vecindad de $\mathbf{x}_0$. Sea $T_N(\mathbf{x} - \mathbf{x}_0)$ el desarrollo de Taylor de grado N de f alrededor de $\mathbf{x}_0$; es decir,

$$T_N(\mathbf{x} - \mathbf{x}_0) = f(\mathbf{x}_0) + d_{\mathbf{x}_0}f(\mathbf{x} - \mathbf{x}_0) + \ldots + \frac{1}{N!}d_{\mathbf{x}_0}^N f(\mathbf{x} - \mathbf{x}_0).$$

Entonces

$$\lim_{\mathbf{x}\to\mathbf{x}_0}\frac{(f(\mathbf{x}) - T_N(\mathbf{x} - \mathbf{x}_0))}{|\mathbf{x} - \mathbf{x}_0|^N} = 0,$$

y T_N es el único polinomio de grado N con esta propiedad. La unicidad dada por el teorema de Taylor muestra que un polinomio de grado N es igual a su desarrollo de Taylor alrededor de un punto arbitrario $\mathbf{x}_0$.

Ejemplo 5. El polinomio $x^2y + x^3 + y^3$ puede ser expresado como un polinomio en $(x - 1)$ y $(y + 1)$ calculando su desarrollo de Taylor en $(1, -1)$. El resultado es

$$x^2y + x^3 + y^3 = -1 + \frac{1}{1!}((x - 1) + 4(y + 1))$$
$$+ \frac{1}{2!}(4(x - 1)^2 + 4(x - 1)(y + 1) - 6(y + 1)^2)$$
$$+ \frac{1}{3!}(6(x - 1)^3 + 6(x - 1)^2(y + 1) + 6(y + 1)^3)$$
$$= -1 + (x - 1) + 4(y + 1) + 2(x - 1)^2$$
$$+ 2(x - 1)(y + 1) - 3(y + 1)^2 + (x - 1)^3$$
$$+ (x - 1)^2(y + 1) + (y + 1)^3.$$

Ejemplo 6. El desarrollo en serie infinita $e^t = 1 + t + (1/2!)t^2 + \ldots$ es válido para todo t. Tomando $t = x + y$, tenemos

$$e^{x+y} = 1 + \frac{1}{1!}(x + y) + \frac{1}{2!}(x^2 + 2xy + y^2) + \ldots, \quad \text{para todo } x \text{ y } y.$$

Por tanto,

$$1 + \frac{1}{1!}(x + y) + \frac{1}{2!}(x^2 + 2xy + y^2)$$

es de desarrollo de Taylor de segundo grado de e^{x+y}. El resto, $(1/3!)(x + y)^2 + \ldots$ tiende a cero si se lo divide por $(\sqrt{x^2 + y^2})^2$ y (x, y) tiende a $(0, 0)$. De acuerdo con el teorema de Taylor, existe un solo polinomio de segundo grado con esta propiedad.

Ejemplo 7. Sea $f(x, y) = e^{xy} \operatorname{sen}(x + y)$. Entonces, como

$$e^{xy} = 1 + xy + \frac{1}{2!}x^2y^2 + R_1$$

y

$$\operatorname{sen}(x + y) = (x + y) - \frac{1}{3!}(x + y)^3 + R_2,$$

podemos multiplicar los dos desarrollos poniendo en el resto todos los términos de grado superior a tres. El resultado es

$$f(x, y) = e^{xy}\operatorname{sen}(x + y) = (x + y) + x^2y + xy^2 - \frac{1}{3!}(x + y)^3 + R,$$

donde $R/|(x, y)|^3$ tiende a cero cuando (x, y) tiende a $(0, 0)$. En otras palabras, hemos encontrado el desarrollo de Taylor de tercer grado de $e^{xy} \operatorname{sen}(x + y)$ alrededor de $(0, 0)$. En su forma estándar, el desarrollo es

$$f(x, y) = e^{xy}\operatorname{sen}(x + y)$$

$$= \frac{1}{1!}(x + y) + \frac{1}{3!}(-x^3 + 3x^2y + 3xy^2 - y^3) + R.$$

Podemos deducir que

$$\frac{\partial^3 f}{\partial y^3}(0, 0) = \frac{\partial^3 f}{\partial x^3}(0, 0) = -1,$$

$$\frac{\partial^3 f}{\partial x\, \partial y^2}(0, 0) = \frac{\partial^3 f}{\partial x^2\, \partial y}(0, 0) = 1.$$

Ejemplo 8. Las funciones e^x y $\cos x$ tienen los siguientes desarrollos de segundo grado alrededor de $x = 0$:

$$e^x = 1 + x + \frac{x^2}{2} + R(x), \qquad \cos x = 1 - \frac{x^2}{2} + R'(x).$$

Por tanto,

$$e^{\cos x} = 1 + \left(1 - \frac{x^2}{2} + R'(x)\right) + \frac{1}{2}\left(1 - \frac{x^2}{2} + R'(x)\right)^2 + R\left(1 - \frac{x^2}{2} + R'(x)\right).$$

Como $R(1 - x^2/2 + R'(x))$ no tiende a cero cuando x tiende a cero, debemos proceder de otra manera para obtener el desarrollo de Taylor de $e^{\cos x}$. Tenemos que

$$e^{\cos x} = e(e^{-1+\cos x})$$

$$= e\left[1 + \left(-\frac{x^2}{2} + R'(x)\right) + \frac{1}{2}\left(-\frac{x^2}{2} + R'(x)\right)^2 + R\left(-\frac{x^2}{2} + R'(x)\right)\right]$$

$$= e\left(1 - \frac{x^2}{2}\right) + R''(x),$$

donde $R''(x)/x^2$ tiende a cero cuando x tiende a cero. También es posible encontrar los coeficientes calculando las derivadas de $e^{\cos x}$.

Ejemplo 9. Como $1/(1 + x^2) = 1 - x^2 + R(x)$,

$$\prod_{k=1}^{n} \frac{1}{1 + x_k^2} = \prod_{k=1}^{n} (1 - x_k^2 + R(x_k))$$

$$= 1 - (x_1^2 + x_2^2 + \ldots + x_n^2) + R'(x_1, \ldots, x_n),$$

donde $R'/|(x_1, \ldots, x_n)|^2$ tiende a cero cuando $(x_1, \ldots, x_n)$ tiende a cero.

Ejemplo 10. El desarrollo de Taylor de $f(\mathbf{x}) = (x_1 + \ldots + x_n)^N$ alrededor de $\mathbf{x}_0 = 0$ es

$$(x_1 + \ldots + x_n)^N = \frac{1}{N!}\left(x_1\frac{\partial}{\partial x_1} + \ldots + x_n\frac{\partial}{\partial x_n}\right)_0^N f$$

$$= \frac{1}{N!}\sum_{k_1+\ldots+k_n=N}\binom{N}{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n} \frac{\partial^N f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}$$

$$= \sum_{k_1+\ldots+k_n=N}\binom{N}{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n}.$$

Solo la diferencial de orden N es distinta de cero en $\mathbf{x}_0 = 0$. Para calcular el coeficiente multinomial, aplicamos el operador $\partial^N/(\partial x_1^{k_1} \ldots \partial x_n^{k_n})$ a ambos lados de la ecuación. Entonces, haciendo $\mathbf{x} = 0$, obtenemos

$$N! = \binom{N}{k_1 \ldots k_n} k_1! \ldots k_n!,$$

de donde se puede obtener la fórmula del Ejemplo 3.

EJERCICIOS

1. Encuentre el desarrollo de Taylor de tercer grado de $(u + v)^3$,
   (a) alrededor del punto $(u_0, v_0) = (0, 0)$,
   (b) alrededor del punto $(u_0, v_0) = (1, 2)$.
2. Encuentre la mejor aproximación de segundo grado de la función $f(x, y) = xe^y$ cerca del punto $(x_0, y_0) = (2, 0)$.
   [*Resp.* $2 + (x - 2) + 2y + y^2 + (x - 2)y$.]
3. Encuentre la mejor aproximación de segundo grado de la función $f(x, y) = x^{y+1}$ cerca del punto $(x_0, y_0) = (2, 0)$.
4. Encuentre los términos cuadráticos del desarrollo de Taylor de xe^{x+y} alrededor de $(0, 0)$.

(a) calculando las derivadas,
(b) por sustitución.

5. Encuentre los términos cuadráticos del desarrollo de Taylor de $e^{\operatorname{sen}(x+y)}$ cerca de (0, 0).

6. Si $f(x, y) = (x^2 + y^2)e^{x^2+y^2}$, use el desarrollo de Taylor de f para calcular

$$\frac{\partial^3 f}{\partial x^2 \partial y}(0, 0).$$

[*Resp*: 0.]

7. Calcule el desarrollo de Taylor de segundo grado de $\sqrt{1 + x^2 + y^2}$ alrededor del punto $(x_0, y_0) = (-1, 1)$.

8. Calcule el desarrollo de Taylor de segundo grado de

(a) $f(x, y, z) = (x^2 + 2xy + y^2)e^z$ cerca de $(x_0, y_0, z_0) = (1, 2, 0)$.
(b) $g(x, y, z) = xy^2z^3$ cerca de $(x_0, y_0, z_0) = (1, 2, -1)$.

9. Exprese el polinomio xy^2z^3 como un polinomio en $(x - 1)$, y y $(z + 1)$.

10. Calcule el desarrollo de Taylor de segundo grado de $\log_2 x$ en $x = 1$. Dibuje el gráfico del desarrollo cerca de $x = 1$.

11. Calcule el desarrollo de Taylor de segundo grado de $\log \cos (x + y)$ en $(x_0, y_0) = (0, 0)$,

(a) calculando las derivadas,
(b) por sustitución.

12. Calcule el desarrollo de Taylor de segundo grado de $\exp(-x_1^2 - x_2^2 - \ldots - x_n^2)$ alrededor de $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$.

13. Demuestre que el desarrollo de Taylor de grado N de un polinomio de grado K, $K \geq N$ está formada por los términos de grado menor o igual que N del polinomio.

14. Calcule las diferencias $d^k_{\mathbf{x}_0} f(\mathbf{y})$ para y arbitrario donde

(a) $k = 2$, $\mathbf{x}_0 = (1, 2)$, $f(x, y) = x^3y + 3x^2 + 2xy^3$.
[*Resp*: $18x^2 + 54xy + 24y^2$.]
(b) $k = 1$, $\mathbf{x}_0 = (a, b, c)$, $f(x, y, z) = 1/(x + y + z + 1)$.
(c) $k = 2$, $\mathbf{x}_0 = (0, 0, 0)$, $f(x, y, z) = 1/(x + y + z + 1)$.
[*Resp*: $2(x + y + z)^2$.]
(d) $k = 4$, $\mathbf{x}_0 = (0, 0, 0)$, $f(x, y, z) = x^3 + 3xy^2 + 4xy^3 + 6x^2y^3 + 7y^5$.
[*Resp*: $96xy^3$.]

15. Encuentre la segunda diferencial de $\operatorname{sen}(x_1 + x_2 + \ldots + x_n)$ en

$$(x_1, x_2, \ldots, x_n) = (0, 0, \ldots 0).$$

4. Desarrollo de Taylor y valores extremos

La tangente $\mathcal{T}$ al gráfico de la función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ en el punto $(\mathbf{x}_0, f(\mathbf{x}_0))$ se encuentra calculando el desarrollo de Taylor de f alrededor de $\mathbf{x}_0$. Consideramos ahora la posibilidad de que el gráfico de f corte a $\mathcal{T}$ en $(\mathbf{x}_0, f(\mathbf{x}_0))$. Las posibilidades en el caso de una variable están indicadas en la Fig. 13.

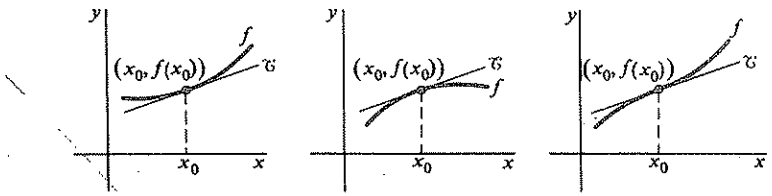


Figura 13

Consideremos primero una situación muy simplificada en la cual f es igual a su desarrollo de Taylor de segundo grado. Es decir, suponemos que

$$f(\mathbf{x}) = f(\mathbf{x}_0) + d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + \tfrac{1}{2} d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0).$$

Los primeros dos términos del desarrollo forman la mejor aproximación afín de f cerca de $\mathbf{x}_0$ y el gráfico de $f(\mathbf{x}_0) + d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es la tangente $\mathcal{T}$ del gráfico de f en $\mathbf{p}_0 = (\mathbf{x}_0, f(\mathbf{x}_0))$. Es claro que si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es positivo para todo $\mathbf{x}$ en $\mathcal{R}^n$ distinto de $\mathbf{x}_0$, entonces

$$f(\mathbf{x}) > f(\mathbf{x}_0) + d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0), \quad \text{para} \quad \mathbf{x} \neq \mathbf{x}_0,$$

lo cual implica que el gráfico de f está por arriba de la tangente $\mathcal{T}$. Análogamente, si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es negativo, salvo para $\mathbf{x} = \mathbf{x}_0$, el gráfico de f está por debajo de $\mathcal{T}$. Por otra parte, si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ cambia de signo en $\mathbf{x}_0$, entonces el gráfico de f cruzará $\mathcal{T}$ en $\mathbf{p}_0$.

Decir que el polinomio cuadrático $Q(\mathbf{x} - \mathbf{x}_0) = d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ cambia de signo en $\mathbf{x}_0$ significa que existen puntos $\mathbf{x}_1$ y $\mathbf{x}_2$ arbitrariamente cercanos a $\mathbf{x}_0$ tales que

$$Q(\mathbf{x}_1 - \mathbf{x}_0) > 0 \quad \text{y} \quad Q(\mathbf{x}_2 - \mathbf{x}_0) < 0. \tag{1}$$

La expresión «arbitrariamente cercano» puede ser omitida, ya que todo polinomio cuadrático homogéneo Q tiene la propiedad

$$Q(t\mathbf{x}) = t^2 Q(\mathbf{x}). \tag{2}$$

Por tanto, si Q es válido para dos vectores $\mathbf{x}_1$ y $\mathbf{x}_2$, no necesariamente cercanos a $\mathbf{x}_0$, entonces (1) es válido también para $t(\mathbf{x}_1 - \mathbf{x}_0) + \mathbf{x}_0$ y $t(\mathbf{x}_2 - \mathbf{x}_0) + \mathbf{x}_0$, para todo $t \neq 0$. Si elegimos t suficientemente pequeño, estos vectores se pueden acercar a $\mathbf{x}_0$ tanto como se desee.

Ejemplo 1. La función

$$f(x, y) = 2x^2 - xy - 3y^2 - 3x + 7y$$

es igual a su desarrollo de Taylor de segundo grado. Tiene un punto crítico en $\mathbf{x}_0 = (1, 1)$ y el plano tangente $\mathcal{T}$ en (1, 1, 2) es horizontal. La segunda diferencial está dada por

$$d^2_{\mathbf{x}_0} f\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} = 4(x-1)^2 - 2(x-1)(y-1) - 6(y-1)^2.$$

Si probamos con $\mathbf{x}_1 = (2, 1)$ y $\mathbf{x}_2 = (1, 2)$, obtenemos

$$d^2_{\mathbf{x}_0} f(\mathbf{x}_1 - \mathbf{x}_0) = 4 > 0,$$
$$d^2_{\mathbf{x}_0} f(\mathbf{x}_2 - \mathbf{x}_0) = -6 < 0.$$

Concluimos que el gráfico de f cruza al plano tangente $\mathcal{T}$ en $(1, 1, 2)$ y, por tanto, que f no tiene un máximo o mínimo local en $\mathbf{x}_0$.

La hipótesis de que f es igual a su desarrollo de Taylor de segundo grado es demasiado restrictiva para que tenga valor práctico. Sin embargo, el teorema siguiente muestra que, bajo hipótesis mucho más generales, el signo de la segunda diferencial todavía determina si f cruza o no a su tangente. El polinomio cuadrático $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es siempre cero para $\mathbf{x} = \mathbf{x}_0$. Si él es positivo para todos los otros puntos, se dice que él es **definido positivo** y si es negativo en los puntos distintos de $\mathbf{x}_0$, se dice que él es **definido negativo**.

4.1 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ una función con derivadas parciales de segundo orden continuas en una vecindad de $\mathbf{x}_0$. Denotemos por $\mathcal{T}$ a la tangente al gráfico de f en el punto $\mathbf{p}_0 = (\mathbf{x}_0, f(\mathbf{x}_0))$. Entonces

(a) Si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es definida positiva, f está por encima de $\mathcal{T}$ en alguna vecindad de $\mathbf{x}_0$.

(b) Si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ es definida negativa, f está por debajo de $\mathcal{T}$ en alguna vecindad de $\mathbf{x}_0$

(c) Si $d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$ toma valores positivos y negativos, f cruza a $\mathcal{T}$ en $\mathbf{p}_0$.

Nótese que no todos los casos posibles están cubiertos por (a), (b) y (c). Por ejemplo, puede suceder que la segunda diferencial sea cero para algún punto distinto de $\mathbf{x}_0$ sin que cambie de signo.

Demostración (de 4.1). Por el teorema de Taylor, tenemos que

$$f(\mathbf{x}) - f(\mathbf{x}_0) - d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) = \tfrac{1}{2} d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + R, \tag{3}$$

donde

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{R}{|\mathbf{x} - \mathbf{x}_0|^2} = 0.$$

Si suponemos (a), debemos demostrar que

$$\tfrac{1}{2} d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + R > 0$$

en alguna vecindad de $\mathbf{x}_0$ y con $\mathbf{x} \neq \mathbf{x}_0$. La homogeneidad del polinomio cuadrático $d^2_{\mathbf{x}_0} f$ [véase Ec. (2)] implica que

$$\frac{d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^2} = d^2_{\mathbf{x}_0} f\left(\frac{\mathbf{x} - \mathbf{x}_0}{|\mathbf{x} - \mathbf{x}_0|}\right).$$

Como $d^2_{\mathbf{x}_0} f$ es definida positiva, sus valores para vectores unitarios tienen una cota inferior $m > 0$ (véase Ejercicio 14 de la Sec. 2). Elegimos $\delta > 0$ tal que, para $0 < |\mathbf{x} - \mathbf{x}_0| < \delta$,

$$\frac{|R|}{|\mathbf{x} - \mathbf{x}_0|^2} \leq \frac{m}{4}.$$

Entonces

$$\tfrac{1}{2} d^2_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + R \geq \frac{m}{2} |\mathbf{x} - \mathbf{x}_0|^2 - \frac{m}{4} |\mathbf{x} - \mathbf{x}_0|^2 > 0$$

que, de acuerdo con la Ec. (3), es la que queríamos demostrar.

La demostración de (b) es prácticamente igual a la recién dada. Para demostrar (c), supongamos que $d^2_{\mathbf{x}_0} f(\mathbf{x}_1 - \mathbf{x}_0) > 0$ y $d^2_{\mathbf{x}_0} f(\mathbf{x}_2 - \mathbf{x}_0) < 0$. Tomemos

$$\mathbf{x}_i(t) = t(\mathbf{x}_i - \mathbf{x}_0) + \mathbf{x}_0, \qquad i = 1, 2, \quad -\infty < t < \infty.$$

Usando la homogeneidad de $d^2_{\mathbf{x}_0} f$ y de la norma, tenemos que si $t \neq 0$, entonces

$$f(\mathbf{x}_i(t)) - f(\mathbf{x}_0) - d_{\mathbf{x}_0} f(\mathbf{x}_i(t) - \mathbf{x}_0)$$
$$= t^2 \left[\tfrac{1}{2} d^2_{\mathbf{x}_0} f(\mathbf{x}_i - \mathbf{x}_0) + |\mathbf{x}_i - \mathbf{x}_0|^2 \frac{R}{|\mathbf{x}_i(t) - \mathbf{x}_0|^2} \right]. \tag{4}$$

Como

$$\lim_{t \to 0} \frac{R}{|\mathbf{x}_i(t) - \mathbf{x}_0|^2} = 0,$$

tenemos que, para todo $t \neq 0$ suficientemente pequeño, el término de la izquierda de la Ec. (4) es positivo si $i = 1$ y negativo si $i = 2$. En otras palabras, el gráfico de f está por encima y por debajo de la tangente $\mathcal{T}$ para valores de $\mathbf{x}$ arbitrariamente cercanos a $\mathbf{x}_0$. Esto termina la demostración.

Ejemplo 2. La función

$$f(x, y) = (x^2 + y^2) e^{x^2 - y^2}$$

tiene puntos críticos en $(0, 0)$, $(0, 1)$ y $(0, -1)$. Esto implica que las tangentes en esos puntos son planos horizontales. El desarrollo de Taylor de segundo grado en estos puntos son

$$f(0, 0) + \tfrac{1}{2} d^2_{(0,0)} f(x, y) = x^2 + y^2,$$

$$f(0, 1) + \tfrac{1}{2} d^2_{(0,1)} f(x, y - 1) = \frac{1}{e} [1 + 2x^2 - 2(y - 1)^2],$$

$$f(0, -1) + \tfrac{1}{2} d^2_{(0,-1)} f(x, y + 1) = \frac{1}{e} [1 + 2x^2 - 2(y + 1)^2].$$

Es claro que $d^2_{(0,0)}f$ es definido positivo, mientras que $d^2_{(0,1)}f$ y $d^2_{(0,-1)}f$ toman valores positivo y negativo arbitrariamente cerca de sus respectivos puntos críticos. Concluimos que f tiene un mínimo local en $(0, 0)$ y no tiene un máximo ni un mínimo local en los puntos $(0, 1)$ y $(0, -1)$.

La Fig. 14 muestra el gráfico de f y los planos tangentes horizontales en $(0, 0, 0)$ y $(0, 1, 1/e)$.

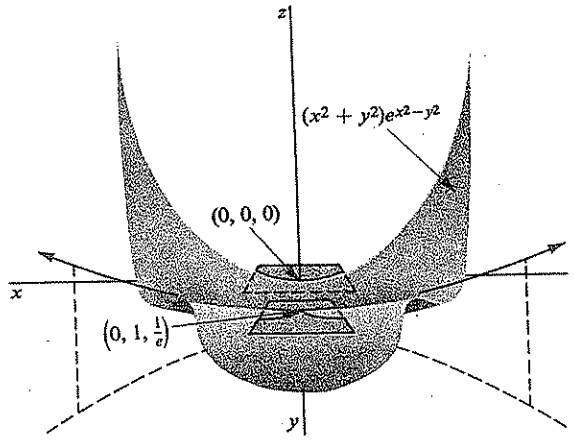


Figura 14

Ejemplo 3. La función

$$f(x, y, z) = x \operatorname{sen} z + z \operatorname{sen} y$$

tiene un punto crítico en $\mathbf{x}_0 = (-1, \pi/2, 0)$. La segunda diferencial en $\mathbf{x}_0$ es

$$d^2_{\mathbf{x}_0} f\left(x + 1, y - \frac{\pi}{2}, z\right) = 2z(x + 1).$$

Como $d^2_{\mathbf{x}_0}f$ toma valores positivos y negativos, la función no tiene ni un máximo ni un mínimo local en $\mathbf{x}_0$.

Ejemplo 4. Si $f(x) = x \operatorname{sen}^3 x$, los cuatro primeros términos del desarrollo de Taylor en $x = 0$ son cero, mientras que

$$d^4_0 fx = 24x^4.$$

Los criterios del Teorema 4.1 no son aplicables a este caso. Sin embargo, una demostración análoga mostraría que f se comporta como su desarrollo de Taylor en cuanto a los cruces de su tangente. La conclusión es que $x \operatorname{sen}^3 x$ tiene un mínimo local en $x = 0$.

No es necesario hacer un examen detallado del polinomio $d^2_{\mathbf{x}_0}f(\mathbf{x} - \mathbf{x}_0)$ para decidir si un punto crítico $\mathbf{x}_0$ de f es un máximo o mínimo local. Es suficiente saber si esta aproximación cuadrática es definida positiva o negativa. Además de los criterios de la Sec. 2, damos a continuación un criterio sencillo para un polinomio cuadrático en dos o tres variables. (Véase Sec. 2, Ejemplo 4 y Ejercicio 16.)

El polinomio

$$ax^2 + 2bxy + cy^2 = (x \quad y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

es definido positivo si, y solo si,

$$a > 0 \quad \text{y} \quad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0$$

y es definido negativo si, y solo si,

$$a < 0 \quad \text{y} \quad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0.$$

El carácter del polinomio

$$ax^2 + by^2 + cz^2 + 2dyz + 2exz + 2fxy = (x \quad y \quad z)\begin{pmatrix} a & f & e \\ f & b & d \\ e & d & c \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

depende del signo de los tres determinantes

$$a, \quad \begin{vmatrix} a & f \\ f & b \end{vmatrix}, \quad \begin{vmatrix} a & f & e \\ f & b & d \\ e & d & c \end{vmatrix}.$$

El polinomio es definido positivo si, y solo si, los tres determinantes son positivos y es definido negativo si, y solo si, el del medio es negativo y los otros dos son positivos.

Ejemplo 5. La función $g(x, y, z) = x^2 + y^2 + z^2 + xy$ tiene puntos críticos solo sí

$$g_x = 2x + y = 0,$$
$$g_y = 2y + x = 0,$$
$$g_z = 2z \quad\quad = 0;$$

de modo que el único punto crítico es $(0, 0, 0)$. Como $d^2_{\mathbf{x}}g = g$, podemos probar si g es positivo definido. Tenemos que

$$a = 1, \quad \begin{vmatrix} a & f \\ f & b \end{vmatrix} = \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{vmatrix} = \frac{3}{4}, \quad \begin{vmatrix} a & f & e \\ f & b & d \\ e & d & c \end{vmatrix} = \begin{vmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{3}{4}.$$

de modo que g es definido positivo y g tiene el valor mínimo 0 en $(0, 0, 0)$.

EJERCICIOS

1. Encuentre todos los puntos críticos de las siguientes funciones:

   (a) $(x + y)e^{-xy}$. [*Resp*: $(\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$.]

   (b) $xy + xz$. [*Resp*: $(0, y, -y)$, todo y.]

   (c) $(x^2 + y^2) \ln (x^2 + y^2)$.

   (d) $\cos (x^2 + y^2 + z^2)$.

   (e) $x^2 + y^2 + z^2$.

2. Calcule el desarrollo de Taylor de segundo grado de la función del Ejercicio 1(a) en cada uno de sus puntos críticos.

3. Clasifique los puntos críticos de 1(a), 1(b) y 1(c) en máximos, mínimos y ninguno de éstos.

4. Para cada una de las funciones siguientes considere la tangente al gráfico de la función en el punto dado. Decida si la función está por encima o por debajo de la tangente cerca del punto indicado o si la cruza en ese punto.

   (a) $x^2 \operatorname{sen} x$ en $x = 1$. [*Resp*. Por encima.]

   (b) $1/(x - y)$ en $(x, y) = (2, 1)$.

   (c) $x^4 + y^4$ en $(0, 0)$.

   (d) $e^{x+w} - x^2 - y^2$ en $(0, 0, 0, 0)$. [*Resp*. Corta.]

5. Encuentre los puntos críticos $\mathbf{x}_0$ de cada una de las funciones siguientes y, examinando a $d^2_{\mathbf{x}_0} f$, decida si la función tiene un máximo local, un mínimo local o ninguno de los anteriores en $\mathbf{x}_0$. Si el examen de la segunda diferencial no da información, considere el término siguiente del desarrollo de Taylor que proporcione información.

   (a) $\operatorname{sen} x \cos x$.

   (b) x^2y^2. [*Resp*: $x = 0$ o $y = 0$, min.]

   (c) $x^2 + 4xy - y^2 - 8x - 6y$.

   (d) $x^2 - xy - y^2 + 5y - 1$. [*Resp*: (1, 2), ninguno]

   (e) $x^2 + 2y^2 - x$.

   (f) $x \operatorname{sen} y$.

   (g) $x^4 + y^4$. [*Resp*: (0, 0), min.]

   (h) $(x - y)^4$.

   (i) $\exp(-x_1^2 - x_2^2 - \ldots - x_n^2)$.

5. Series de Fourier

El desarrollo de Taylor de grado N de una función f requiere la existencia de por lo menos las derivadas N-ésimas en el punto x_0. Además, como las derivadas $f^{(k)}(x_0)$ que aparecen en el desarrollo de Taylor están determinadas por las propiedades de f en una vecindad de x_0, el desarrollo de Taylor solo aproxima a f en una vecindad de x_0. Otro método para atacar el problema de la aproximación de una función es el de usar un **polinomio trigonométrico** de grado N.

$$s_N(x) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos kx + b_k \operatorname{sen} kx). \tag{1}$$

En vez de usar los valores cerca de un punto en el dominio de la función f que se desea aproximar, el método de Fourier permite calcular los coeficientes a_k y b_k tomando ciertos términos promedios de f en el intervalo en que ella está definida. Primero supondremos que f es continua en el intervalo $-\pi \le x \le \pi$ y definimos los **coeficientes de Fourier** de f por las fórmulas

$$\mathbf{5.1} \qquad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx, \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} kx \, dx.$$

El polinomio trigonométrico que se obtiene al sustituir estos valores de a_k y y b_k en la Ec. (1) se puede usar a menudo para aproximar f. [De hecho, en el Ejercicio 2(b) se muestra que si f es un polinomio trigonométrico, entonces esta sustitución en (1) da f.] Los resultados de los Ejemplos 1 y 2 que siguen, dan una idea de lo buena que es esta elección de coeficientes. (Véase también el Teorema 5.2.)

El polinomio trigonométrico (1), con los coeficientes determinados por las Ecs. 5.1, se llama la **aproximación N-ésima de Fourier** de f en el intervalo $[-\pi, \pi]$. Es claro que el polinomio es una función periódica de periodo 2π, es decir, satisface $s_N(x + 2\pi) = s_N(x)$ para todo x real. De la periodicidad se tiene que

$$s_N(x + 2k\pi) = s_N(x)$$

para todo x y todos los enteros k. Por esta razón, es posible restringirse a los x en un intervalo fijo de largo 2π, digamos, $-\pi \le x \le \pi$. Calcularemos algunos ejemplos de aproximaciones de Fourier para tener una idea de cómo funcionan.

Ejemplo 1. Sea $f(x) = |x|$ para $-\pi \le x \le \pi$. Entonces

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \cos kx \, dx, \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \operatorname{sen} kx \, dx.$$

Es claro que $|x| \operatorname{sen} kx$ tiene integral cero en $[-\pi, \pi]$, ya que las integrales en $[-\pi, 0]$ y $[0, \pi]$ son negativas una de la otra. Por tanto, $b_k = 0$ para $k = 1, 2, \ldots$ Por otra parte, el gráfico de $|x| \cos kx$ es simétrico con respecto del eje y, de modo que, para $k \neq 0$, podemos integrar por partes, obteniendo

$$\begin{aligned} a_k &= \frac{2}{\pi} \int_0^{\pi} x \cos kx \, dx \\ &= \frac{2}{\pi} \left[\frac{x \operatorname{sen} kx}{k} \right]_0^{\pi} - \frac{2}{\pi k} \int_0^{\pi} \operatorname{sen} kx \, dx \\ &= \left[\frac{2}{\pi k^2} \cos kx \right]_0^{\pi} = \frac{2}{\pi k^2} (\cos k\pi - 1) \\ &= \begin{cases} 0, & k = 2, 4, 6, \ldots, \\ -\dfrac{4}{\pi k^2}, & k = 1, 3, 5, \ldots. \end{cases} \end{aligned}$$

Si $k = 0$, tenemos

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x\,dx = \pi.$$

Resumiendo,

$$a_0 = \pi, \qquad a_k = \begin{cases} 0, & k = 2, 4, 6, \ldots, \\ -\dfrac{4}{\pi k^2}, & k = 1, 3, 5, \ldots, \end{cases}$$

$$b_k = 0, \qquad k = 1, 2, 3, \ldots.$$

Por tanto, la aproximación N-ésima de Fourier está dada, para $N = 1, 3, 5, \ldots$, por el polinomio trigonométrico

$$s_N(x) = \frac{\pi}{2} - \frac{4}{\pi}\cos x - \frac{4}{\pi}\frac{\cos 3x}{3^2} - \cdots - \frac{4}{\pi}\frac{\cos Nx}{N^2}.$$

Si N es par, tenemos que $s_N(x) = s_{N-1}(x)$. La Fig. 15 muestra cómo los gráficos de s_0, s_1 y s_3 aproximan al gráfico de $|x|$ en $[-\pi, \pi]$.

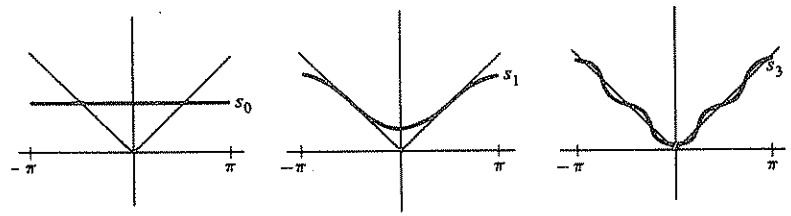


Figura 15

Ejemplo 2. Sea

$$g(x) = \begin{cases} 1, & 0 \le x \le \pi \\ -1, & -\pi \le x < 0. \end{cases}$$

Entonces los coeficientes de Fourier de g se obtienen por integración y son

$$a_k = 0, \qquad k = 0, 1, 2, \ldots$$

$$b_k = \begin{cases} 0, & k = 2, 4, 6, \ldots \\ \dfrac{4}{\pi k}, & k = 1, 3, 5, \ldots. \end{cases}$$

Por tanto, si N es impar, la aproximación N-ésima de Fourier de g es

$$s_N(x) = \frac{4}{\pi}\operatorname{sen} x + \frac{4}{\pi}\frac{\operatorname{sen} 3x}{3} + \ldots + \frac{4}{\pi}\frac{\operatorname{sen} Nx}{N}.$$

La Fig. 16 muestra los gráficos de s_1, s_3 y s_5 en relación con el de g. Es importante saber si, para valores específicos de x, la aproximación de Fourier $s_N(x)$ converge a $f(x)$ cuando $N \to \infty$, donde f es la función usada para calcular los coeficientes de Fourier. Por definición, la **serie de Fourier** de f es la serie

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \operatorname{sen} kx), \tag{2}$$

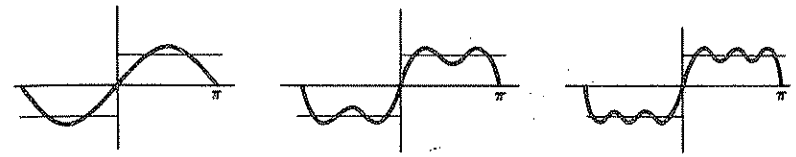


Figura 16

donde a_k y b_k están dados por la Ec. 5.1. El Teorema 5.2 da condiciones bajo las cuales la serie de Fourier de f puede ser usada para representar f. En efecto, supongamos que el gráfico de f es **liso por pedazos.** Esto significa que el intervalo $(-\pi, \pi)$ puede ser dividido en subintervalos de extremos $-\pi < x_1 < x_2 < \ldots < x_k < \pi$, de modo que f pueda ser extendida del intervalo (x_k, x_{k+1}) a una función continuamente diferenciable en el intervalo $[x_k, x_{k+1}]$. Entonces se puede demostrar que la serie de Fourier de f converge a $f(x)$ en los puntos donde f es continua y en un punto de discontinuidad x_k converge al valor promedio $\frac{1}{2}(f(x_k-) + f(x_k+))$, donde $f(x-)$ es el límite izquierdo de f en x y $f(x+)$ el límite derecho. La Fig. 17 muestra el gráfico de una función lisa por pedazos donde los valores medios están indicados por un punto.

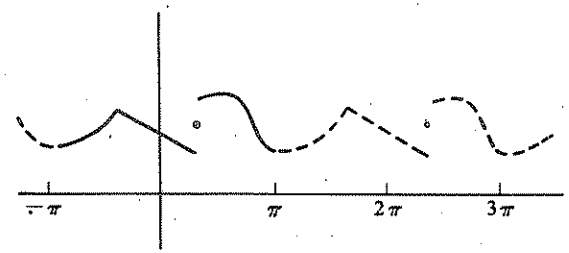


Figura 17

5.2 Teorema

Sea f una función lisa por pedazos en $[-\pi, \pi]$. Entonces la serie de Fourier de f converge en todo punto x del intervalo al valor $(\frac{1}{2})[f(x-) + f(x+)]$. En particular, si f es continua en x, la serie converge a $f(x)$. En $x = \pm\pi$, la serie converge a $(\frac{1}{2})[f(\pi-) + f(-\pi+)]$.

La demostración será dada en el Apéndice.

Los Ejemplos 1 y 2 dan una idea de cómo convergen las sumas parciales de una serie de Fourier. En cada uno de estos ejemplos, la función es lisa por pedazos, por tanto, la serie converge al valor apropiado de la función para $-\pi < x < \pi$.

Ejemplo 3. La función g del Ejemplo 2 ha sido definida arbitrariamente, de modo que tome el valor 1 en $x = 0$. A pesar de esto, el Teorema 5.2 nos permite concluir que la serie de Fourier de g converge como sigue:

$$\sum_{k=0}^{\infty} \frac{4}{\pi} \frac{\operatorname{sen}(2k+1)x}{2k+1} = \begin{cases} 1, & 0 < x < \pi, \\ 0, & x = 0, \\ -1, & -\pi < x < 0. \end{cases}$$

Más específicamente, podemos tomar $x = \pi/2$, obteniendo el desarrollo en serie alternante

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4}.$$

El Teorema 5.2 muestra la razón para elegir los coeficientes en el polinomio trigonométrico de acuerdo con la Ec. 5.1. Esto se debe a que, bajo condiciones favorables, la sucesión de polinomios trigonométricos que se obtiene converge a la función f. Ahora daremos otra razón. Consideremos el espacio vectorial $\mathcal{T}_N$ de los polinomios trigonométricos de grado N, restringidos al intervalo $-\pi \le x \le \pi$. Es claro que las $2N - 1$ funciones

$$\frac{1}{\sqrt{2}}, \cos x, \operatorname{sen} x, \ldots, \cos Nx, \operatorname{sen} Nx \tag{3}$$

generan el espacio vectorial $\mathcal{T}_N$. Por tanto, la dimensión de $\mathcal{T}_N$ es a lo más $2N - 1$. Si en $\mathcal{T}_N$ introducimos el producto interno

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) g(x)\, dx \tag{4}$$

podemos comprobar que el conjunto (3) de funciones es un conjunto ortonormal y, de acuerdo con el Teorema 8.1 del Cap. 2, es linealmente independiente. Tenemos el siguiente teorema.

5.3 Teorema

El conjunto (3) de funciones satisface las condiciones de ortogonalidad

$$\langle \cos km, \cos nx \rangle = \begin{cases} 1, & m = n \neq 0, \\ 0, & m \neq n, \end{cases}$$

$$\langle \operatorname{sen} mx, \operatorname{sen} nx \rangle = \begin{cases} 1, & m = n, \\ 0, & m \neq n, \end{cases}$$

$$\langle \cos mx, \operatorname{sen} nx \rangle = 0, \quad \text{todo } m, n,$$

$$\left\langle \frac{1}{\sqrt{2}}, \cos mx \right\rangle = \left\langle \frac{1}{\sqrt{2}}, \operatorname{sen} mx \right\rangle = 0, \quad \text{todo } m,$$

$$\left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle = 1.$$

La demostración es un cálculo rutinario de integrales y se deja como ejercicio. (Por ejemplo, la relación)

$$\langle \cos mx, \operatorname{sen} nx \rangle = 0$$

es equivalente a

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos mx \operatorname{sen} nx\, dx = 0.)$$

Del Teorema 5.3, se deduce que las funciones trigonométricas (3) forman una base ortonormal de $\mathcal{T}_N$. Además, del Teorema 8.1 del Cap. 2, se deduce que los coeficientes en la expresión

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^{N} (a_k \cos kx + b_k \operatorname{sen} kx)$$

se obtienen mediante las fórmulas

$$a_0 = \left\langle \frac{1}{\sqrt{2}}, f(x) \right\rangle,$$

$$a_k = \langle \cos kx, f(x) \rangle, \qquad k > 0,$$

$$b_k = \langle \operatorname{sen} kx, f(x) \rangle, \qquad k > 0.$$

Por la definición del producto interno, estas fórmulas son las de 5.1, excepto que en 5.1 la fórmula para a_0 tiene el factor $1/\sqrt{2}$ incluido en el polinomio trigonométrico. Por tanto, las fórmulas para los coeficientes de Fourier son las adecuadas para calcular los coeficientes respecto de la base ortonormal (3) y el producto interno (4). Esto explica por qué no es necesario volver a calcular los coeficientes encontrados para el espacio $\mathcal{T}_N$ cuando se considera $\mathcal{T}_{N+1}$, incluyendo $\cos (N+1)x$ y $\operatorname{sen} (N+1)x$.

Terminamos esta sección mostrando cómo se puede simplificar algunas de las fórmulas sobre series de Fourier, usando funciones exponenciales complejas. Recordemos que en la Sec. 7 del Cap. 2 definimos

$$e^{\pm ikx} = \cos kx \pm i \operatorname{sen} kx.$$

Las identidades

$$\cos kx = \frac{e^{ikx} + e^{-ikx}}{2}, \qquad \operatorname{sen} kx = \frac{e^{ikx} - e^{-ikx}}{2i} \tag{5}$$

se obtienen inmediatamente por sustitución. Por tanto, el polinomio trigonométrico (1) puede ser escrito en la forma

$$s_N(x) = \sum_{k=-N}^{N} c_k e^{ikx}, \tag{6}$$

donde

$$c_k = \frac{a_k - ib_k}{2}, \qquad k > 0,$$

$$c_0 = \frac{a_0}{2}, \tag{7}$$

$$c_k = \frac{a_{-k} + ib_{-k}}{2}, \qquad k < 0.$$

Por tanto, nuevamente por sustitución, para todo entero k,

5.4 $$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} f(x)\,dx.$$

La Ec. 5.4 es la versión compleja de 5.1. Evidentemente, en general, los coeficientes c_k son números complejos. Pero si f es real, el polinomio trigonométrico (6) también lo es y es igual al polinomio (1), siempre que los coeficientes satisfagan las relaciones (6). Las ventajas del uso de la forma compleja quedan en evidencia al considerar la relación más sencilla de ortogonalidad obtenida usando el producto interno complejo

$$\langle f, g\rangle_c = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\bar{g}(x)\,dx.$$

En términos de este producto interno tenemos

$$\langle e^{ikx}, e^{ilx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx}e^{-ilx}\,dx$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-l)x}\,dx = \begin{cases}1, & k = l,\\ 0, & k \neq l.\end{cases}$$

Los detalles de estos cálculos se dejan como ejercicio.

EJERCICIOS.

1. Calcule el coeficiente de Fourier de cada una de las funciones siguientes. Dibuje el gráfico de cada función y de las tres primeras aproximaciones de Fourier $s_N(x)$ que sean distintas.

   (a) $f_1(x) = x, \qquad -\pi \le x \le \pi.$ [*Resp*: $b_k = (2(-1)^{k+1})/k$.]

   (b) $f_2(x) = \begin{cases}(-\pi - x), & -\pi \le x \le 0.\\ (\pi - x), & 0 < x \le \pi.\end{cases}$

   (c) $f_3(x) = x^2, \qquad -\pi \le x \le \pi.$

2. (a) Definiendo el producto interno $\langle f, g\rangle$ de dos funciones continuas por

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx,$$

demuestre las relaciones de ortogonalidad

$$\langle \cos nx, \cos mx\rangle = \begin{cases}1, & m = n \neq 0\\ 0, & m \neq n\end{cases}$$

$$\langle \operatorname{sen} nx, \operatorname{sen} mx\rangle = \begin{cases}1, & m = n \neq 0\\ 0, & m \neq n\end{cases}$$

$$\langle \operatorname{sen} nx, \cos mx\rangle = 0, \text{ todo } n, m,$$

donde n y m son enteros.

(b) Use el resultado de la parte (a) para demostrar que si

$$f(x) = \frac{c_0}{2} + \sum_{k=1}^{N}(c_k \cos kx + d_k \operatorname{sen} kx),$$

entonces los coeficientes de Fourier de f están dados por

$$\langle f(x), \cos kx\rangle = c_k, \qquad k = 0, 1, 2, \ldots, N$$

y

$$\langle f(x), \operatorname{sen} kx\rangle = d_k, \qquad k = 1, 2, \ldots, N,$$

y que $\langle f(x), \cos kx\rangle = \langle f(x), \operatorname{sen} kx\rangle = 0$ si $k > N$.

3. Use las identidades de la Ec. (5) para expresar los siguientes polinomios trigonométricos como polinomios (es decir, combinaciones lineales de potencias) en e^{ikx} y e^{-ikx}.

   (a) $1 + \cos x$.

   (b) $\cos x - 1 \operatorname{sen} 2x$.

4. El resultado del Ejercicio 2(b) muestra que si una función f puede ser expresada como un polinomio trigonométrico, entonces ese polinomio es el desarrollo de Fourier de f. Por ejemplo, la identidad

$$\cos^2 x = \tfrac{1}{2} + \tfrac{1}{2}\cos 2x$$

expresa $\cos^2 x$ como un polinomio trigonométrico, de modo que ese es el desarrollo de Fourier de $\cos^2 x$. Encuentre el desarrollo de Fourier de las siguientes funciones:

   (a) $\operatorname{sen}^2 x$.

   (b) $\cos^3 x$.

   (c) $\operatorname{sen} 2x \cos x$.

5. (a) Demuestre la identidad

$$\frac{1}{2} + \sum_{k=1}^{N}\cos ku = \frac{\operatorname{sen}(N + 1/2)u}{2\operatorname{sen}(u/2)}.$$

[*Indicación*: Sume para k de 1 a N la identidad $2 \operatorname{sen}(u/2)\cos ku = \operatorname{sen}(k + \frac{1}{2})u$.]

   (b) ¿Es la suma de la izquierda el desarrollo de Fourier de la función de la derecha?

6. Sea f una función compleja definida en $[-\pi, \pi]$ y defina $c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} f(x)\,dx$, $k = 0, \pm 1, \pm 2, \ldots$ Entonces $\sum_{k=-N}^{N} c_k e^{ikx}$ se llama la N-ésima aproximación compleja de Fourier de f.

(a) Demuestre que si

$$f(x) = \sum_{k=-N}^{N} d_k e^{-ikx},$$

entonces las constantes d_k son los coeficientes complejos de Fourier de f.

(b) Demuestre que si f es real, entonces

$$2c_0 = a_0, \qquad 2c_k = a_k - ib_k,$$

y

$$2c_{-k} = a_k + ib_k \text{ para } k = 1, 2, 3, \ldots.$$

7. Demuestre que si s_N y t_N son las aproximaciones N-ésimas de Fourier de f y g, respectivamente, entonces $as_N + bt_N$ es la aproximación N-ésima de Fourier de $af + bg$, donde a y b son constantes.

6. Desarrollo de Fourier modificado

La aplicación directa de los métodos de Fourier a los problemas prácticos generalmente debe ser modificada con respecto a lo hecho en la sección anterior. En esta sección describiremos estas modificaciones y calcularemos algunos ejemplos.

A pesar de que el intervalo $[-\pi, \pi]$ es natural para los desarrollos de Fourier, puesto que éste es el intervalo de un periodo de las funciones trigonométricas, puede suceder que se desee aproximar una función en otro intervalo.

Si la función f que se desea aproximar está definida en el intervalo $[-p, p]$, se puede hacer un cambio en los cálculos como sigue. Definimos

$$f_p(x) = f\left(\frac{px}{\pi}\right), \qquad -\pi \le x \le \pi.$$

Entonces podemos calcular los coeficientes de Fourier de f_p usando la Ec. 5.1. Los polinomios trigonométricos s_N obtenidos con ellos, aproximan f_p en $[-\pi, \pi]$. Para aproximar f en $[-p, p]$, consideremos

$$s_N\left(\frac{\pi x}{p}\right) = \frac{a_0}{2} + \sum_{k=1}^{N}\left(a_k \cos\frac{k\pi x}{p} + b_k \operatorname{sen}\frac{k\pi x}{p}\right), \qquad -p \le x \le p$$

Los coeficientes a_k y b_k pueden ser calculados directamente en términos de f haciendo un cambio de variables. Tenemos que

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f_p(x)\cos kx\,dx = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{px}{\pi}\right)\cos kx\,dx$$

$$= \frac{1}{p}\int_{-p}^{p} f(x)\cos\left(\frac{k\pi x}{p}\right)dx.$$

b_k se calcula en forma análoga y tenemos las fórmulas

6.1 $$a_k = \frac{1}{p}\int_{-p}^{p} f(x)\cos\frac{k\pi x}{p}\,dx, \qquad b_k = \frac{1}{p}\int_{-p}^{p} f(x)\operatorname{sen}\frac{k\pi x}{p}\,dx$$

para los coeficientes de la aproximación de Fourier

$$\frac{a_0}{2} + \sum_{k=1}^{N}\left(a_k \cos\frac{k\pi x}{p} + b_k \operatorname{sen}\frac{k\pi x}{p}\right)$$

de la función f definida en $[-p, p]$.

Ejemplo 1. Si

$$h(x) = \begin{cases} 1, & 0 \le x \le p \\ -1, & -p \le x \le 0, \end{cases}$$

entonces

$$a_k = 0, \qquad k = 0, 1, 2, \ldots,$$

$$b_k = \frac{2}{p}\int_0^p \operatorname{sen}\frac{k\pi x}{p}\,dx$$

$$= \frac{2}{\pi}\int_0^\pi \operatorname{sen} kx\,dx = \begin{cases} 0, & k = 2, 4, 6, \ldots, \\ \dfrac{4}{\pi k}, & k = 1, 3, 5, \ldots. \end{cases}$$

Por tanto, la aproximación N-ésima de Fourier de h está dada, para N impar, por

$$s_N(x) = \frac{4}{\pi}\operatorname{sen}\frac{\pi x}{p} + \frac{4}{3\pi}\operatorname{sen}\frac{3\pi x}{p} + \ldots + \frac{4}{N\pi}\operatorname{sen}\frac{N\pi x}{p}, \qquad -p \le x \le p.$$

Para una función f definida en un intervalo arbitrario $a \le x \le b$ es útil considerar una extensión periódica f_E de f con periodo $b - a$ y definida para todos los reales x. La Fig. 18 muestra una extensión de este tipo. Tomamos $2p = b - a$, de modo que $p = (b - a)/2$ y $-p = -(b - a)/2$. Calculamos los coeficientes de Fourier de f_E en el intervalo $[-p, p]$ de acuerdo con la Ec. 6.1. Además, como los integrandos en la Ec. 6.1 tienen periodo $2p$, podemos usar el hecho geométricamente obvio de que la integración puede ser hecha en cual-

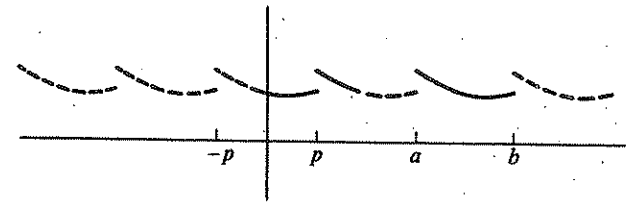


Figura 18

quier intervalo de largo $2p = b - a$, en particular, en $[a, b]$. Por tanto, la Ec. 6.1 puede ser reescrita en la forma

$$\textbf{6.2} \qquad a_k = \frac{2}{b-a}\int_a^b f(x)\cos\frac{2k\pi x}{b-a}\,dx,$$

$$b_k = \frac{2}{b-a}\int_a^b f(x)\operatorname{sen}\frac{2k\pi x}{b-a}\,dx.$$

Los polinomios trigonométricos asociados son

$$s_N(x) = \frac{a_0}{2} + \sum_{k=0}^{N}\left(a_k\cos\frac{2k\pi x}{b-a} + b_k\operatorname{sen}\frac{2k\pi x}{b-a}\right).$$

Ejemplo 2. Sea $f(x) = x$, para $0 \le x \le 1$. Integrando por partes, tenemos que

$$a_k = 2\int_0^1 x\cos 2k\pi x\,dx$$

$$= 2\left[x\frac{\operatorname{sen}2k\pi x}{2k\pi}\right]_0^1 - \frac{2}{2k\pi}\int_0^1 \operatorname{sen}2k\pi x\,dx = 0,$$

$$b_k = 2\int_0^1 x\operatorname{sen}2k\pi x\,dx$$

$$= 2\left[-x\frac{\cos 2k\pi x}{2k\pi}\right]_0^1 + \frac{2}{2k\pi}\int_0^1 \cos 2k\pi x\,dx$$

$$= -\frac{\cos 2k\pi}{k\pi} = -\frac{1}{k\pi}.$$

La serie de Fourier es

$$\frac{1}{2} - \frac{1}{\pi}\left(\operatorname{sen}2\pi x + \frac{\operatorname{sen}4\pi x}{2} + \frac{\operatorname{sen}6\pi x}{3} + \dots\right).$$

A veces, es más conveniente tener un desarrollo usando solo cosenos o solo senos. Empezamos notando que coseno es una función **par**; es decir, $\cos(-x) = \cos xy$ que seno es **impar**, es decir, $\operatorname{sen}(-x) = -\operatorname{sen}x$. En general, una función f es par si $f(-x) = f(x)$ para todo x en el dominio de f y es impar si $f(-x) = -f(x)$ para todo x. La Fig. 19 muestra los gráficos de algunas funciones pares e impares. Geométricamente, una función es par si su gráfico es simétrico con respecto del eje y, y es impar si es simétrico respecto del origen. Por tanto,

$$\int_{-p}^{p} f(x)\,dx = 2\int_0^p f(x)\,dx, \quad \textbf{para todo } f \tag{1}$$

y

$$\int_{-p}^{p} f(x)\,dx = 0, \quad \textbf{para todo } f. \tag{2}$$

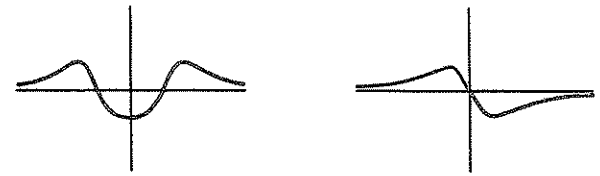


Figura 19

Por tanto, si f es una función par periódica, el producto

$$f(x)\operatorname{sen}\frac{\pi x}{p}$$

es impar. (¿Por qué?) Entonces, por 6.1, el coeficiente de Fourier b_k es

$$b_k = \frac{1}{p}\int_{-p}^{p} f(x)\operatorname{sen}\frac{\pi x}{p}\,dx = 0. \tag{3}$$

Por tanto, el desarrollo de Fourier de una función par solo tiene términos en cosenos. Análogamente, si f es impar periódica, el producto

$$f(x)\cos\frac{\pi x}{p}$$

es también impar, de modo que su coeficiente de Fourier a_k es

$$a_k = \frac{1}{p}\int_{-p}^{p} f(x)\cos\frac{\pi x}{p}\,dx = 0. \tag{4}$$

Luego, *el desarrollo de Fourier de una función impar solo tiene términos en senos.*

Los resultados del párrafo anterior son la base de la solución del siguiente problema: dada una función $f(x)$, definida en el intervalo $0 \le x \le p$, encontrar un desarrollo de f en serie trigonométrica formada solo por términos en coseno o solo en seno. El truco es extender la definición de f del intervalo $0 \le x \le p$ a todos los reales x, de modo tal que esta extensión sea periódica de periodo $2p$ y par o impar. Entonces calculamos la serie de Fourier de la extensión. Si f_e es una extensión par de f, entonces el desarrollo de Fourier de f_e solo contendrá términos en coseno y coincidirá con f en $0 \le x \le p$. Análogamente, si f_o es una extensión periódica impar de f, entonces el desarrollo de Fourier de f_o solo contiene términos en seno y coincidirá con f en $0 \le x \le p$. Ilustramos este procedimiento mediante un par de ejemplos.

Ejemplo 3. Calcularemos el desarrollo en cosenos de la función definida por $f(x) = 1 - x$ para $0 \le x \le 2$. Consideramos la extensión periódica par mostrada en la Fig. 20. Esta extensión f_e está dada definiendo $f_e(x) = f(-x)$ para $-2 \le x \le 0$ y luego extendiendo periódicamente con periodo 4 a todo

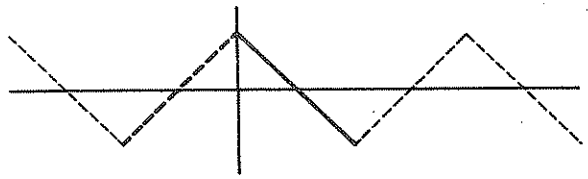


Figura 20

el eje x. Podemos usar la Ec. 6.1 para calcular los coeficientes de Fourier de f_e. Como f_e es par, la Ec. (3) muestra que $b_k = 0$ para todo k. Además, la Ec. (1) permite escribir

$$a_k = \frac{1}{2}\int_{-2}^{2} f_e(x)\cos\frac{k\pi x}{2}\,dx$$

$$= \int_{0}^{2} f_e(x)\cos\frac{k\pi x}{2}\,dx.$$

Como para $0 \le x \le 2$ f_e coincide con $f(x) = 1 - x$, tenemos que para $k > 0$

$$a_k = \int_0^2 (1-x)\cos\frac{k\pi x}{2}\,dx$$

$$= \left[\frac{2}{k\pi}(1-x)\operatorname{sen}\frac{k\pi x}{2}\right]_0^2 + \frac{2}{k\pi}\int_0^2 \operatorname{sen}\frac{k\pi x}{2}\,dx$$

$$= \frac{4}{\pi^2 k^2}[1 - \cos k\pi]$$

$$= \begin{cases} 0, & k \text{ par}, \\ \dfrac{8}{\pi^2 k^2}, & k \text{ impar}, \end{cases}$$

Por último,

$$a_0 = \int_0^2 (1-x)\,dx = 0.$$

Por tanto, el desarrollo en cosenos de f en $0 \le x \le 2$ tiene término general

$$\frac{8}{\pi^2 k^2}\cos\frac{k\pi x}{2}, \quad k \text{ par},.$$

El desarrollo tiene la forma

$$\frac{2}{\pi^2}\left(\cos\pi x + \frac{\cos 2\pi x}{4} + \frac{\cos 3\pi x}{9} + \cdots\right).$$

Ejemplo 4. Consideremos la misma función del Ejemplo 3, $f(x) = 1 - x$ para $0 \le x \le 2$ y calculemos su desarrollo en términos de senos usando la

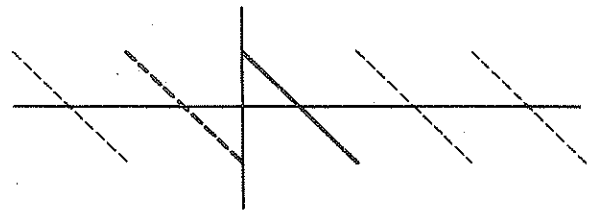


Figura 21

función impar periódica de la Fig. 21. Primero definimos $f_o(x) = -f(-x)$ para $-2 \le x \le 0$ y luego extendemos periódicamente con periodo 4. Usando la Ec. 6.1 y la Ec. (4) encontramos que, como queríamos, $a_k = 0$ para todo k. Además, por la Ec. (1),

$$b_k = \frac{1}{2}\int_{-2}^{2} f_o(x)\operatorname{sen}\frac{k\pi x}{2}\,dx$$

$$= \int_0^2 f_o(x)\operatorname{sen}\frac{k\pi x}{2}\,dx.$$

Pero $f_o(x) = 1 - x$ para $0 \le x \le 2$, de modo que

$$b_k = \int_0^2 (1-x)\operatorname{sen}\frac{k\pi x}{2}\,dx$$

$$= \left[-\frac{2}{k\pi}(1-x)\cos\frac{k\pi x}{2}\right]_0^2 - \frac{1}{k\pi}\int_0^2 \cos\frac{k\pi x}{2}\,dx$$

$$= \left[\frac{2}{k\pi}(-1)^k + \frac{2}{k\pi}\right] - \frac{1}{k\pi}\left[\frac{2}{k\pi}\operatorname{sen}\frac{k\pi x}{2}\right]_0^2$$

$$= \begin{cases} 0, & k \text{ impar}. \\ \dfrac{4}{k\pi}, & k \text{ par}, \end{cases}$$

Por tanto, el término general en el desarrollo en senos es

$$\frac{4}{k\pi}\operatorname{sen}\frac{k\pi x}{2}, \quad \text{par. } k.$$

y el desarrollo en senos es

$$\frac{2}{\pi}\left(\operatorname{sen}\pi x + \frac{\operatorname{sen} 2\pi x}{2} + \frac{\operatorname{sen} 3\pi x}{3} + \cdots\right).$$

EJERCICIOS

1. Encuentre la serie de Fourier de la función

$$f(x) = -x, \qquad -2 < x < 2.$$

¿A qué valores converge esta serie para $x = 2$ y $x = -2$?

2. Encuentre la serie de Fourier de la función

$$f(x) = 1 + x, \qquad 1 < x < 2.$$

¿A qué valores converge esta serie para $x = 1$ y $x = 2$?

3. Encuentre (a) el desarrollo de Fourier en coseno y (b) en seno de la función

$$f(x) = x, \qquad 0 < x < \pi.$$

(c) Compare los resultados de (a) y (b) con el desarrollo de Fourier de

$$g(x) = x, \qquad -\pi < x < \pi.$$

4. Demuestre que toda función real f definida en un intervalo simétrico $[-a, a]$ puede ser escrita como la suma de una función par f_e y una función impar f_o. [*Indicación*: Sea $f_e(x) = (f(x) + f(-x))/2$.]

5. Demuestre que si se definen las combinaciones apropiadas,

   (a) el producto de funciones pares es par,
   (b) el producto de funciones impares es par,
   (c) el producto de una función par y una función impar es impar,
   (d) una combinación lineal de funciones pares es par y una combinación lineal de funciones impares es impar.

6. Use las propiedades elementales de la integral para demostrar las Ecs. (1) y (2).

7. Sea una función impar en $[-\pi, \pi]$ y sea g una función par. Sea a_k, b_k y a'_k, b'_k los coeficientes de Fourier de f y g, respectivamente. Demuestre que

$$a_k = 0, \qquad b_k = \frac{2}{\pi}\int_0^\pi f(x) \operatorname{sen} kx\, dx,$$

$$a'_k = \frac{2}{\pi}\int_0^\pi g(x) \cos kx\, dx, \qquad b'_k = 0.$$

7. Las ecuaciones del calor y de las ondas

En esta sección mostraremos cómo las series de Fourier pueden ser usadas para resolver problemas de conducción de calor y de movimientos de ondas. Primero encontramos una ecuación diferencial que sea satisfecha por la magnitud física que se estudia y luego aplicamos las series de Fourier para resolver la ecuación.

Supongamos que tenemos un alambre delgado de densidad uniforme y largo p. Sea $u(x, t)$ su temperatura, en el instante t, en el punto situado a x unidades de un extremo. Por tanto, $0 \le x \le p$ y podemos suponer $t \ge 0$. Supondremos que el único traspaso de calor es a lo largo del alambre y que la temperatura en los dos extremos se mantiene fija. Por esta razón, sin pérdida de generalidad, podemos representar el alambre por un segmento de recta a lo largo del eje x y considerar la temperatura como el gráfico de una función $u = u(x, t)$, tal como en el ejemplo mostrado en la Fig. 22.

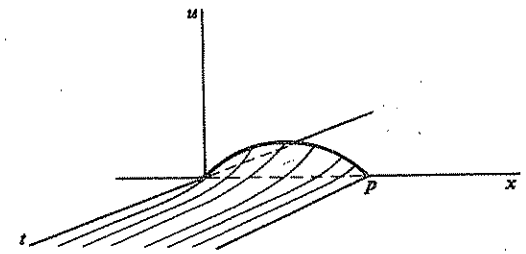


Figura 22

El principio básico de la conducción de calor es que el flujo calórico es proporcional a, y en la dirección del gradiente de temperatura ∇u. Recordemos que ∇u es la dirección en la cual la temperatura crece con máxima rapidez, de modo que es razonable pensar que el calor fluye en la dirección opuesta, de más caliente a más frío. Como el medio es de una dimensión y es representado por un segmento del eje x, el gradiente está representado por $\partial u/\partial x$. Por tanto, la velocidad de cambio total de calor en el segmento $[x_1, x_2]$ es proporcional a

$$-\frac{\partial u}{\partial x}(x_1, t) + \frac{\partial u}{\partial x}(x_2, t). \tag{1}$$

Pero la velocidad de cambio de calor es también proporcional al cambio promedio de la temperatura del intervalo,

$$\int_{x_1}^{x_2} \frac{\partial u}{\partial t}(x, t)\, dx, \tag{2}$$

Por el teorema fundamental del cálculo, la expresión (1) puede ser escrita en la forma

$$\int_{x_1}^{x_2} \frac{\partial^2 u}{\partial x^2}(x, t)\, dx. \tag{3}$$

Por tanto, las expresiones proporcionales (2) y (3) para la velocidad de cambio total de calor pueden ser combinadas, obteniéndose

$$a^2 \int_{x_1}^{x_2} \frac{\partial^2 u}{\partial x^2}(x, t)\, dx = \int_{x_1}^{x_2} \frac{\partial u}{\partial t}(x, t)\, dx,$$

donde a^2 es una constante positiva de proporcionalidad. Si permitimos que x_2 varíe, podemos diferenciar ambos términos de la ecuación con respecto de x_2, obteniendo

$$a^2 \frac{\partial^2 u}{\partial x^2}(x, t) = \frac{\partial u}{\partial t}(x, t). \tag{7.1}$$

Esta es la **ecuación del calor** o **de difusión** en una dimensión.

La Ec. 7.1 es lineal en el sentido de que si u_1 y u_2 son soluciones, entonces también lo es toda combinación lineal $b_1u_1 + b_2u_2$. Para especificar soluciones particulares, imponemos condiciones lineales de frontera de la forma

$$u(0, t) = 0 \quad \text{y} \quad u(p, t) = 0, \tag{4}$$

y una condición inicial de la forma

$$y(x, 0) = h(x).$$

El método estándar de resolución es el de **separación de las variables**, con el cual tratamos de encontrar soluciones productos de la forma

$$u(x, t) = G(x)H(t)$$

con condiciones de frontera $G(0) = G(p) = 0$. Si existe una solución de este tipo, sustituyendo en $a^2u_{xx} = u_t$, se tiene

$$a^2G''(x)H(t) = G(x)H'(t),$$

para $0 \leq x \leq p$, $0 < t$. Dividiendo por $G(x)H(t)$, se tiene

$$a^2 \frac{G''(x)}{G(x)} = \frac{H'(t)}{H(t)}. \tag{5}$$

Para que esta ecuación sea válida para todo x y t, ambos lados deben ser constantes y denotamos por $-\lambda^2$ al valor de ambos lados de la ecuación. Este procedimiento es el origen del término separación de las variables.

Haciendo iguales a $-\lambda^2$ ambos lados de la Ec. (5), se tiene que

$$a^2G'' + \lambda^2G = 0, \tag{6}$$

$$H' + \lambda^2H = 0. \tag{7}$$

La Ec. (6) tiene las soluciones

$$G(x) = c_1 \cos\left(\frac{\lambda}{a}\right)x + c_2 \operatorname{sen}\left(\frac{\lambda}{a}\right)x.$$

pero $G(0) = 0$ requiere que $c_1 = 0$ y $G(p) = 0$ que $c_2 \operatorname{sen}(-\lambda/a)p = 0$. Esta condición se puede lograr sin hacer $c_2 = 0$ solo si elegimos λ de modo que $(\lambda/a)p = k\pi$, donde k es un entero. Es decir, debemos tomar $\lambda = (ka\pi)/p$ de modo que G tiene la forma

$$G(x) = c_2 \operatorname{sen}\left(\frac{k\pi}{p}\right)x, \qquad k = 1, 2, \ldots.$$

La Ec. (7) tiene la solución

$$H(t) = e^{-\lambda^2 t}$$

que, ya que $\lambda^2 = (k^2a^2\pi^2)/p^2$, es

$$H(t) = e^{-k^2(a^2\pi^2/p^2)t}.$$

La solución producto $u(x, t)$ es, salvo un factor constante,

$$u_k(x, t) = e^{-k^2(a^2\pi^2/p^2)t} \operatorname{sen}\left(\frac{k\pi}{p}\right)x, \qquad k = 1, 2, \ldots.$$

Por tanto, un límite de combinaciones lineales

$$\sum_{k=1}^{N} b_k u_k(x, t)$$

tiene la forma

$$u(x, t) = \sum_{k=1}^{\infty} b_k e^{-k^2(a^2\pi^2/p^2)t} \operatorname{sen}\left(\frac{k\pi}{p}\right)x. \tag{8}$$

Para satisfacer la condición inicial $u(x, 0) = h(x)$, exigimos

$$\sum_{k=1}^{\infty} b_k \operatorname{sen}\left(\frac{k\pi}{p}\right)x = h(x), \tag{9}$$

para algún N. Si h es de esta forma, la Ec. (8) da una solución del problema. Las condiciones frontera $u(0, t) = u(p, t) = 0$ requieren que la temperatura en los extremos sea cero y la condición inicial $u(x, 0) = h(x)$ especifica la temperatura inicial en cada punto del alambre. Por esta razón, surge naturalmente el problema de encontrar un desarrollo de Fourier en senos para $h(x)$. En la Sec. 8 consideramos las condiciones bajo las cuales la serie (8) es una solución de la Ec. 7.1.

Ejemplo 1. Para ser más específicos acerca de la solución de la ecuación de difusión de las ondas, suponemos por simplicidad que $p = \pi$ y recordamos que para resolver la ecuación $a^2u_{xx} = u_t$ con la condición de frontera $u(0, t) = u(\pi, t) = 0$ y la condición inicial $u(x, 0) = h(x)$, en general queremos poder expresar h como una serie infinita de la forma

$$h(x) = \sum_{k=1}^{\infty} b_k \operatorname{sen} kx. \tag{10}$$

Supongamos, por ejemplo, que h está dada en $[0, \pi]$ por

$$h(x) = \begin{cases} x, & 0 \leq x \leq \pi/2 \\ \pi - x, & \pi/2 \leq x \leq \pi. \end{cases}$$

La Fig. 23 muestra el gráfico de h.

Para que la Ec. (10) represente el desarrollo de Fourier de h en $[0, \pi]$, exten-

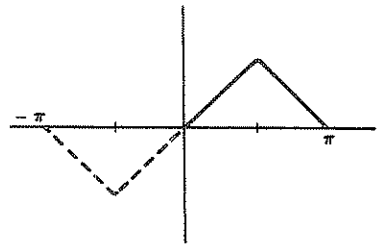


Figura 23

demos h al intervalo $[-\pi, \pi]$ de modo que los términos que involucran al coseno en el desarrollo de h sean cero, de modo que solo sea necesario calcular los términos que involucran al seno. Para ello, extendemos h simétricamente con respecto del origen, como lo muestra la Fig. 23. Entonces

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} h(x)\cos kx\,dx = 0,$$

ya que $h(-x)\cos k(-x) = -h(x)\cos kx$ las integrales en $[0, \pi]$ y $[-\pi, 0]$ son negativas una de la otra. Para calcular b_k usamos el hecho de que $h(-x)\operatorname{sen} k(-x) = h(x)\operatorname{sen} kx$, de modo que el gráfico de esta función es simétrico con respecto del eje y. Entonces,

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} h(x)\operatorname{sen} kx\,dx \\ &= \frac{2}{\pi}\int_{0}^{\pi} h(x)\operatorname{sen} kx\,dx \\ &= \frac{2}{\pi}\int_{0}^{\pi/2} x\operatorname{sen} kx\,dx + \frac{2}{\pi}\int_{\pi/2}^{\pi}(\pi - x)\operatorname{sen} kx\,dx \\ &= \frac{4}{\pi k^2}\operatorname{sen}\left(\frac{k\pi}{2}\right). \end{aligned}$$

Por tanto,

$$b_k = \begin{cases} 0, & k \text{ par.} \\ \dfrac{4}{\pi k^2}, & k = 1, 5, 9, \ldots, \\ \dfrac{-4}{\pi k^2}, & k = 3, 7, 11, \ldots. \end{cases}$$

El Teorema 5.2 implica que

$$h(x) = \frac{4}{\pi}\left(\frac{\operatorname{sen} x}{1^2} - \frac{\operatorname{sen} 3x}{3^2} + \frac{\operatorname{sen} 5x}{5^2} - \frac{\operatorname{sen} 7x}{7^2} + - \ldots\right)$$

para todo x en $[0, \pi]$. Finalmente, de la Ec. (8) se espera que la solución de la ecuación $a^2u_{xx} = u_t$ esté dada por

$$u(x,t) = \frac{4}{\pi}\left(e^{-a^2t}\operatorname{sen} x - \frac{1}{3^2}e^{-3^2a^2t}\operatorname{sen} 3x + \frac{1}{5^2}e^{-5^2a^2t}\operatorname{sen} 5x - \frac{1}{7^2}e^{-7^2a^2t}\operatorname{sen} 7x + - \ldots\right)$$

La verificación de que $u(0, t) = u(\pi, t)$ es evidente si hacemos $x = 0$ y $x = \pi$. Si hacemos $t = 0$ obtenemos la representación de h por su serie de Fourier que está garantizada por el Teorema 5.2. La verificación de que $u(x, t)$ satisface la ecuación $a^2u_{xx} = u_t$ se hace efectuando la derivación término a término de la serie de u. Esto se justificará en la próxima sección a pesar de que en el Ejercicio 7 se deja la verificación formal.

A continuación consideramos la ecuación de ondas en una dimensión. Como motivación física, consideremos una cuerda elástica tensa de largo p y densidad uniforme ρ colocada en el eje x de $\mathcal{R}^3$. Supongamos que los extremos están fijos en $x = 0$ y $x = p$ mediante fuerzas opuestas de magnitud F. Si se hace vibrar la cuerda, nuestro problema es determinar la posición $\mathbf{x}(s, t)$ en $\mathcal{R}^3$ en el instante t del punto de la cuerda que está a una distancia s del extremo fijo en $x = 0$. La Fig. 24 muestra una configuración posible. Imaginemos que la cuerda está subdividida en pedazos cortos de largo Δs y obtengamos dos expresiones distintas para la fuerza total que actúa en un segmento típico de

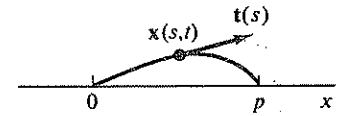


Figura 24

la subdivisión. Si $\mathbf{t}(s)$ es el vector unitario tangente a la cuerda en $\mathbf{x}(s)$, entonces, las fuerzas opuestas que actúan en $\mathbf{x}(s_0)$ y $\mathbf{x}(s_0 + \Delta s)$ son

$$F\mathbf{t}(x_0 + \Delta s) \quad \text{y} \quad -F\mathbf{t}(s_0).$$

Por tanto, la fuerza total es

$$F[\mathbf{t}(s_0 + \Delta s) - \mathbf{t}(s_0)].$$

Por otra parte, por la ley de Newton, la fuerza es igual a la masa, $\rho\,\Delta s$, multiplicada por la aceleración $\mathbf{a}(s_0)$. Por tanto,

$$\rho\mathbf{a} = F\left[\frac{\mathbf{t}(s_0 + \Delta s) - \mathbf{t}(s_0)}{\Delta s}\right].$$

Pero

$$\mathbf{a} = \frac{\partial^2\mathbf{x}}{\partial t^2}(s, t) \quad \text{y} \quad \mathbf{t}(s) = \frac{\partial\mathbf{x}}{\partial s}(s, t);$$

de modo que, si $\Delta s \to 0$, tenemos

$$\rho \frac{\partial^2 \mathbf{x}}{\partial t^2}(s, t) = F \frac{\partial^2 \mathbf{x}}{\partial s^2}(s, t). \tag{11}$$

Esta ecuación diferencial vectorial es equivalente al sistema de tres ecuaciones escalares

$$\frac{\partial^2 x}{\partial t^2} = a^2 \frac{\partial^2 x}{\partial s^2},$$
$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial s^2},$$
$$\frac{\partial^2 z}{\partial t^2} = a^2 \frac{\partial^2 z}{\partial s^2},$$

donde hemos tomado $a^2 = F/\rho > 0$. Formalmente, el problema de encontrar soluciones $x(s, t)$, $y(s, t)$ y $z(s, t)$ es el mismo. Sin embargo, en la práctica se considera que el movimiento longitudinal (a lo largo del eje x) es pequeño y no se considera la ecuación para $x(s, t)$. Las otras dos ecuaciones tienen el mismo significado a no ser que se hagan suposiciones especiales. Para ser específicos, suponemos que la cuerda ha sido pulsada de modo tal que su movimiento se efectúa en el plano xy, con $z(s, t) \equiv 0$. Finalmente suponemos que los desplazamientos son lo suficientemente pequeños como para que podamos remplazar $\partial^2 y/\partial s^2$ por $\partial^2 y/\partial x^2$. Esta última suposición requiere justificación experimental en los casos prácticos. Por tanto, con alguna pérdida de generalidad, en lugar de la ecuación vectorial (11) consideramos una sola ecuación escalar para $y(x, t)$, a saber,

7.2
$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2}.$$

Esta ecuación diferencial no especifica completamente la vibración de una cuerda a no ser que impongamos algunas condiciones iniciales:

$$y(x, 0) = f(x), \tag{12}$$
$$\frac{\partial y}{\partial t}(x, 0) = g(x). \tag{13}$$

La primera condición especifica el desplazamiento inicial ($t = 0$) en la dirección y como función de x y la segunda ecuación especifica la velocidad inicial en la dirección y también como función de x. Además imponemos las condiciones de frontera

$$y(0, t) = 0 \quad \text{y} \quad y(p, t) = 0. \tag{14}$$

Como en el caso de la ecuación del calor, usamos la separación de variables, usando la linealidad de la Ec. 7.2 y las condiciones (14) para construir soluciones que también satisfagan (12) y (13). Probamos con

$$y(x, t) = G(x)H(t),$$

con lo cual 7.2 toma la forma

$$G(x)H''(t) = a^2 G''(x)H(t)$$

o bien

$$\frac{H''(t)}{H(t)} = a^2 \frac{G''(x)}{G(x)}.$$

Como el lado izquierdo es independiente de t, ambos lados son constantes, de modo que escribimos

$$\frac{G''(x)}{G(x)} = \lambda \quad \text{y} \quad \frac{H''(t)}{H(t)} = a^2\lambda.$$

La primera ecuación, $G''(x) = \lambda G(x)$ tiene soluciones

$$G(x) = c_1 e^{\sqrt{\lambda}\, x} + c_2 e^{-\sqrt{\lambda}\, x}.$$

Las condiciones de frontera (14) requieren que

$$G(0) = 0 \quad \text{y} \quad G(p) = 0;$$

de modo que

$$c_1 + c_2 = 0 \quad \text{y} \quad c_1 e^{\sqrt{\lambda}\, p} + c_2 e^{-\sqrt{\lambda}\, p} = 0.$$

Si despejamos c_1 y c_2 vemos que, para que existan soluciones distintas de cero, debemos tener que $c_1 = -c_2$ y $e^{-\sqrt{\lambda}p} = e^{\sqrt{\lambda}p}$, es decir, $e^{2\sqrt{\lambda}p} = 1$. Por tanto, si aceptamos exponentes complejos, concluimos que $2\sqrt{\lambda}p = 2\pi k i$ para algún entero k, de modo que

$$\sqrt{\lambda} = \frac{\pi k i}{p}.$$

Por tanto, las soluciones $G(x)$ son de la forma

$$G(x) = c_1 e^{(\pi k i/p)x} - c_1 e^{(-\pi k i/p)x}$$
$$= 2c_1 i \operatorname{sen} \frac{\pi k}{p} x.$$

Escribimos

$$G_k(x) = b_k \operatorname{sen} \frac{\pi k}{p} x.$$

La ecuación diferencial $H''(t) = a^2\lambda H(t)$ ahora tiene la forma

$$H''(t) - \frac{\pi^2 k^2 a^2}{p^2} H(t) = 0,$$

puesto que hemos determinado que $\lambda = \pi^2 k^2/p^2$. Las soluciones de esta ecuación son de la forma

$$H_k(t) = C_k \cos \frac{\pi k a}{p} t + D_k \operatorname{sen} \frac{\pi k a}{p} t;$$

por tanto, las soluciones producto $y_k = G_k H_k$ toman la forma

$$y_k(x, t) = \left[A_k \cos \frac{\pi k a}{p} t + B_k \operatorname{sen} \frac{\pi k a}{p} t\right] \operatorname{sen} \frac{\pi k}{p} x.$$

Para satisfacer las condiciones iniciales (12) y (13) formamos sumas finitas o infinitas del tipo

$$y(x, t) = \sum_{k=0}^{\infty} \left[A_k \cos \frac{\pi k a}{p} t + B_k \operatorname{sen} \frac{\pi k a}{p} t\right] \operatorname{sen} \frac{\pi k}{p} x.$$

Por tanto, las condiciones iniciales son formalmente

$$y(x, 0) = \sum_{k=0}^{\infty} A_k \operatorname{sen} \frac{\pi k}{p} x = f(x) \tag{15}$$

y

$$y_t(x, 0) = \sum_{k=0}^{\infty} \frac{\pi k a}{p} B_k \operatorname{sen} \frac{\pi k}{p} x = g(x). \tag{16}$$

Los coeficientes A_k y $(\pi k a/p)B_k$ se determinan de modo que ellos sean los coeficientes de los senos en el desarrollo de Fourier de f y g respectivamente.

Ejemplo 2. Como un ejemplo sencillo, consideramos una cuerda que está inicialmente estacionaria, de modo que $y_t(x, 0) = 0$. Suponemos que el desplazamiento inicial está dado por

$$g(x, 0) = b \operatorname{sen} \frac{\pi x}{p}, \tag{16}$$

y que $y(0, t) = y(p, t) = 0$. Elegimos $B_k = 0$ en la Ec. (16) y $A_1 = b$ en la Ec. (15), con $A_k = 0$ para $k \neq 1$. (En este caso f y g son tan sencillas que sus desarrollos de Fourier son evidentes por simple inspección.) Entonces la solución de 7.2 toma la forma

$$y(x, t) = b \cos \frac{\pi a}{p} t \operatorname{sen} \frac{\pi}{p} x.$$

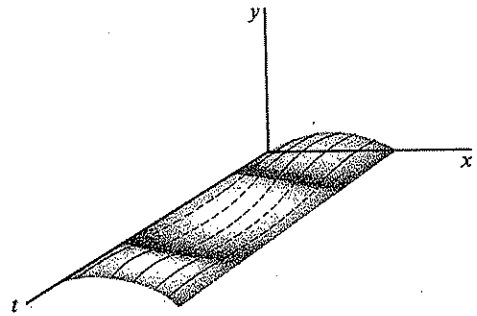


Figura 25

Recordemos que para que la Ec. 7.2 tenga significado físico, es necesario que las vibraciones de la cuerda sean pequeñas, en otras palabras, el coeficiente b debe ser pequeño. La Fig. 25 muestra el gráfico de la solución cuando $a = 1$, $p = \pi$, y b ha sido elegido irrealmente grande para recalcar algunas propiedades cualitativas de la situación.

EJERCICIOS

1. Use el método de separación de variables para resolver la ecuación de calor en una dimensión $a^2 u_{xx} = u_t$ con las condiciones de frontera e inicial siguientes:
   (a) $u(0, t) = u(p, t) = 0$, $u(x, 0) = \operatorname{sen}(\pi x/p)$.
   (b) $u(0, t) = u(\pi, t) = 0$, $u(x, 0) = x(\pi - x)$.
   (c) $u_x(0, t) = u_x(\pi, t) = 0$, $u(x, 0) = \operatorname{sen} x$.

2. Use el método de separación de variables para resolver la ecuación de ondas en una dimensión: $a^2 u_{xx} = u_{tt}$, con las condiciones de frontera e inicial siguientes:
   (a) $u(0, t) = u(\pi, t) = 0$, $u(x, 0) = \operatorname{sen} x$, $u_t(x, 0) = 0$.
   (b) $u(0, t) = u(\pi, t) = 0$, $u(x, 0) = \begin{cases} x, & 0 \leq x \leq \pi/2 \\ \pi - x, & \pi/2 \leq x \leq \pi \end{cases}$, $u_t(x, 0) = 0$.
   (c) $u(0, t) = u(\pi, t) = 0$, $u(x, 0) = 0$, para $0 \leq x \leq \pi$,
   $$u_t(x, 0) = \begin{cases} 0, & 0 \leq x < \pi/2 \\ 1, & \pi/2 \leq x \leq \pi. \end{cases}$$

3. La ecuación de calor en dos dimensiones es $a^2(u_{xx} + u_{yy}) = u_t$ y si u es independiente de t, $u_t \equiv 0$, lo cual conduce a la ecuación de Laplace $u_{xx} + u_{yy} = 0$ para la temperatura de la ecuación de calor en estado estable de flujo en dos dimensiones. Use el método de separación de variables para resolver la ecuación $u_{xx} + u_{yy} = 0$ en el rectángulo $0 \leq x \leq \pi$, $0 \leq y \leq \pi$ con la condición de frontera $u(0, y) = u(\pi, y) = 0$, $u(x, 0) = 0$, $u(x, \pi) = \operatorname{sen} x$.

4. Sea L el operador definido por
   (a) $L(u) = u_t - a^2 u_{xx}$.
   (b) $L(y) = y_{tt} - a^2 y_{xx}$.
   Demuestre que L es un operador lineal y deduzca de ello qué combinaciones lineales de (a) la ecuación del calor y (b) la ecuación de ondas son soluciones.

5. (a) Demuestre que una condición de frontera de la forma $u(a, t) = u(b, t) = 0$ es «lineal» en el sentido de que si dos funciones la satisfacen, entonces toda combinación lineal de ellas también la satisface.
   (b) Demuestre que una condición de frontera de la forma $u(a, t) = 1$ no es lineal en el sentido de la parte (a).
   (c) Demuestre que la condición inicial $u(x, t) = f(x)$ no es lineal en el sentido de la parte (a) a no ser que f sea idénticamente cero.

6. La ecuación de Laplace en dos dimensiones en coordenadas polares es, de acuerdo con el Ejercicio 13 de la Sec. 5 del Cap. 4,

$$r^2\frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0.$$

(a) Haciendo $u(r, \theta) = G(r)H(\theta)$ demuestre que el método de separación de variables conduce a las siguientes ecuaciones diferenciales:

$$H'' - \lambda H = 0,$$

$$r^2G'' + rG' + \lambda G = 0.$$

(b) Demuestre que $H'' - \lambda H = 0$ tiene soluciones que satisfacen $H(0) = H(2\pi)$ si, y solo si, $\lambda = -k^2$, donde k es un entero y que las soluciones son $\cos k\theta$ y $\operatorname{sen} k\theta$.

(c) Demuestre que $r^2G'' + rG' - k^2G = 0$ tiene soluciones r^k y r^{-k} para $k = 0, 1, 2, \ldots$; pero que los exponentes negativos no son posibles si se desea que $u(r, \theta) = G(r)H(\theta)$ sea finito para $r = 0$.

(d) Demuestre que si

$$f(\theta) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos k\theta + b_k \operatorname{sen} k\theta),$$

entonces

$$u(r, \theta) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k r^k \cos k\theta + b_k r^k \operatorname{sen} k\theta)$$

satisface la ecuación de Laplace en coordenadas polares con la condición de frontera $u(1, \theta) = f(\theta)$.

7. Verifique que el desarrollo en series para $u(x, t)$ dada por la Ec. (8) es una solución formal de $a^2u_{xx} = u_t$. Use la diferenciación término a término de la serie. (Esto se justifica en el Teorema 8.5 de la sección que sigue.)

8. Convergencia uniforme

Sea $f_k(\mathbf{x})$, $k = 1, 2, 3, \ldots$ una sucesión de funciones reales definidas para todo $\mathbf{x}$ en algún conjunto S. Entonces para todo $\mathbf{x}$, consideramos la serie $\sum_{k=1}^{\infty} f_k(\mathbf{x})$. Si ella converge para todo $\mathbf{x}$ en S, decimos que la serie **converge puntualmente** en S. Denotamos por $f(\mathbf{x})$ el límite para cada $\mathbf{x}$ en S y tenemos que

$$f(\mathbf{x}) = \sum_{k=1}^{\infty} f_k(\mathbf{x}) = \lim_{N\to\infty} \sum_{k=1}^{N} f_k(\mathbf{x}).$$

Recordemos que esto significa que, para todo $\mathbf{x}$ en S, existe un número $f(\mathbf{x})$ con la propiedad de que, dado $\epsilon > 0$, existe un entero K tal que

$$\left| \sum_{k=1}^{N} f_k(\mathbf{x}) - f(\mathbf{x}) \right| < \epsilon,$$

cuando $N \geq K$.

Ejemplo 1. La serie $\sum_{k=0}^{\infty} x^k$ tiene como suma parcial $(N+1)$-ésima

$$\sum_{k=0}^{N} x_k = \begin{cases} \dfrac{1 - x^{N+1}}{1 - x}, & x \neq 1, \\ N + 1, & x = 1. \end{cases}$$

Entonces

$$\sum_{k=0}^{\infty} x^k = \lim_{N\to\infty} \sum_{k=0}^{N} x^k = \frac{1}{1-x}, \quad \text{para} \quad -1 < x < 1.$$

Para valores de x fuera del intervalo $(-1, 1)$, la serie no converge.

La serie trigonométrica $\sum_{k=1}^{\infty} (\operatorname{sen} kx)/k^2$ converge puntualmente para todo real x. Esto se debe a que sus términos pueden ser comparados con los de la serie convergente $\sum_{k=1}^{\infty} 1/k^2$, notando que

$$\left| \frac{\operatorname{sen} kx}{k^2} \right| \leq \frac{1}{k^2}, \qquad k = 1, 2, \ldots.$$

El resultado es que la serie dada converge absolutamente.

Una serie infinita $\sum_{k=1}^{\infty} f_k(\mathbf{x})$ que converge, para todo $\mathbf{x}$ en un conjunto S, a un número $f(\mathbf{x})$ define una función f en S. Sin embargo, en general, se puede decir muy poco acerca de las propiedades de f si solo se usa la convergencia puntual. Por esta razón, a veces es útil considerar una forma más fuerte de convergencia en S. Decimos que $\sum_{k=1}^{\infty} f_k$ **converge uniformemente** a la función f en S si, dado $\epsilon > 0$, existe un entero K tal que, para todo $\mathbf{x}$ en S y para todo $N \geq K$,

$$\left| \sum_{k=1}^{N} f_k(\mathbf{x}) - f(\mathbf{x}) \right| < \epsilon.$$

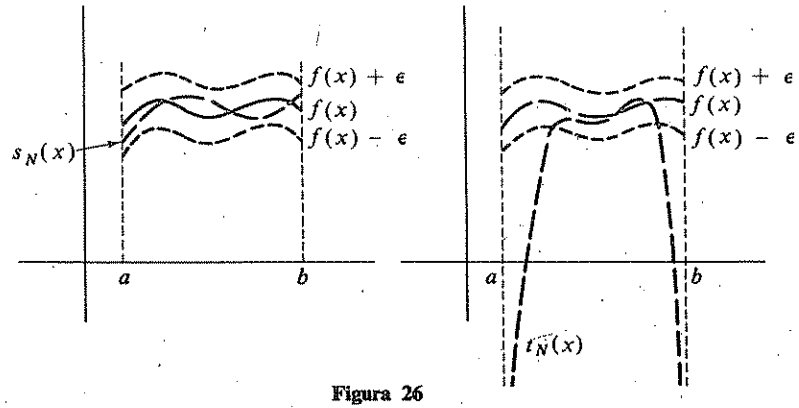


Figura 26

Esta definición debe ser comparada cuidadosamente con la definición de convergencia puntual. Nótese que la convergencia uniforme implica la convergencia puntual, pero la implicación inversa no es verdadera en general. En grandes líneas, la convergencia uniforme de una serie de funciones definidas en un conjunto S significa que la serie converge en cada punto por lo menos con una rapidez determinada. En el caso de convergencia puntual la serie puede tener puntos en los cuales la convergencia es arbitrariamente lenta. La Fig. 26 muestra ejemplos de convergencia uniforme y no uniforme a una misma función f; $s_N(x)$ y $t_N(x)$ son las sumas parciales N-ésimas de las dos series.

Para determinar si una serie converge uniformemente tenemos el siguiente

8.1 Criterio de Weierstrass

Sea $\sum_{k=1}^{\infty} f_k$ una serie de funciones reales definidas en un conjunto S. Si existe una serie de constantes $\sum_{k=1}^{\infty} p_k$, tal que

1. $|f_k(\mathbf{x})| \leq p_k$ para todo $\mathbf{x}$ en S y $k = 1, 2, \ldots,$
2. $\sum_{k=1}^{\infty} p_k$ converge,

entonces $\sum_{k=1}^{\infty} f_k$ converge uniformemente a una función f definida en S.

Demostración. El criterio de comparación para series demuestra que $\sum_{k=1}^{\infty} f_k(\mathbf{x})$ converge (incluso absolutamente) para cada $\mathbf{x}$ en S a un número que denotaremos por $f(\mathbf{x})$. Por tanto, tenemos que

$$f(\mathbf{x}) - \sum_{k=1}^{N} f_k(\mathbf{x}) = \sum_{k=1}^{\infty} f_k(\mathbf{x}) - \sum_{k=1}^{N} f_k(\mathbf{x})$$

$$= \sum_{k=N+1}^{\infty} f_k(\mathbf{x}).$$

De donde

$$\left| f(\mathbf{x}) - \sum_{k=1}^{N} f_k(\mathbf{x}) \right| \leq \sum_{k=N+1}^{\infty} |f_k(\mathbf{x})| \leq \sum_{k=N+1}^{\infty} p_k.$$

Como $\sum_{k=1}^{\infty} p_k$ converge, dado $\epsilon > 0$ podemos encontrar K tal que $\sum_{k=N}^{\infty} p_k$ si $N > K$. Esto completa la demostración, ya que el número K depende solo de ϵ y no de $\mathbf{x}$.

Ejemplo 2. La serie trigonométrica $\sum_{k=1}^{\infty} (\operatorname{sen} kx)/k^2$ converge uniformemente para todo x real, ya que

$$\left| \frac{\operatorname{sen} kx}{k^2} \right| \leq \frac{1}{k^2},$$

y $\sum_{k=1}^{\infty} 1/k^2$ converge. Sin embargo, la serie de potencias $\sum_{k=0}^{\infty} x^k$, a pesar de converger puntualmente para $-1 < x < 1$ no converge uniformemente en $(-1, 1)$. Véase el Ejercicio 6. El criterio de Weierstrass puede ser aplicado en cualquier subintervalo cerrado $[-r, r]$ notando que $|x^k| \leq r^k$ para x en $[-r, r]$ y que $\sum_{k=0}^{\infty} r^k$ converge si $0 \leq r \leq 1$. Por tanto, la serie de potencias converge puntualmente en $(-1, 1)$ y uniformemente en $[-r, r]$ para todo $r < 1$.

Los cuatro teoremas siguientes se refieren a la convergencia uniforme de series de funciones. Ellos afirman que ciertas operaciones de límites pueden ser intercambiadas con la suma de las series siempre que la convergencia sea uniforme. Si la convergencia uniforme se remplaza por la convergencia puntual, las afirmaciones resultantes son falsas. Véase el Ejercicio 9.

8.2 Teorema

Sea $f_1, f_2, f_3, \ldots$ una sucesión de funciones definidas en un subconjunto S de $\mathfrak{R}^n$. Supongamos que $\mathbf{x}_0$ es un punto límite de S y que el límite

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f_k(\mathbf{x})$$

existe para $k = 1, 2, \ldots$ Entonces

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \sum_{k=1}^{\infty} f_k(\mathbf{x}) = \sum_{k=1}^{\infty} \lim_{\mathbf{x} \to \mathbf{x}_0} f_k(\mathbf{x}),$$

siempre que la serie de la derecha converja y la serie de la izquierda converja uniformemente en S.

Demostración. Sea $\lim_{\mathbf{x} \to \mathbf{x}_0} f_k(\mathbf{x}) = a_k$. Entonces, sumando y restando $\sum_{k=1}^{N} f_k(\mathbf{x})$ y $\sum_{k=1}^{N} a_k$, tenemos que

$$\left| \sum_{k=1}^{\infty} f_k(\mathbf{x}) - \sum_{k=1}^{\infty} a_k \right| \leq \left| \sum_{k=1}^{\infty} f_k(\mathbf{x}) - \sum_{k=1}^{N} f_k(\mathbf{x}) \right| + \left| \sum_{k=1}^{N} f_k(\mathbf{x}) - \sum_{k=1}^{N} a_k \right| + \left| \sum_{k=1}^{N} a_k - \sum_{k=1}^{\infty} a_k \right|. \tag{1}$$

Ahora, sea $\epsilon > 0$. Como $\sum_{k=1}^{\infty} f_k$ converge uniformemente, podemos elegir K de modo que $N > K$ implique

$$\left| \sum_{k=1}^{\infty} f_k(\mathbf{x}) - \sum_{k=1}^{N} f_k(\mathbf{x}) \right| < \frac{\epsilon}{3}, \text{ para todo } \mathbf{x} \text{ en } S.$$

Entonces elegimos $N > K$ tal que

$$\left| \sum_{k=1}^{N} a_k - \sum_{k=1}^{\infty} a_k \right| < \frac{\epsilon}{3}.$$

Finalmente, elegimos $\delta > 0$ tal que $|\mathbf{x} - \mathbf{x}_0| < \delta$ implique, usando la relación $\lim_{\mathbf{x}\to\mathbf{x}_0} \sum_{k=1}^{N} f_k(\mathbf{x}) = \sum_{k=1}^{N} a_k$, que

$$\left| \sum_{k=1}^{N} f_k(\mathbf{x}) - \sum_{k=1}^{N} a_k \right| < \frac{\epsilon}{3}.$$

Entonces, para todo x tal que $|\mathbf{x} - \mathbf{x}_0| < \delta$ el lado izquierdo de (1) es menor que ϵ.

8.3 Corolario

Si $\sum_{k=1}^{\infty} f_k$ es una serie uniformemente convergente de funciones continuas f_k definidas en un subconjunto S de $\mathcal{R}^n$, entonces la función f definida por $f(\mathbf{x}) = \sum_{k=1}^{\infty} f_k(\mathbf{x})$ es continua en S.

En los próximos dos teoremas nos restringiremos a funciones de una variable a pesar de que ellos pueden ser aplicados al caso de varias variables considerando las variables una a una.

8.4 Teorema

Si la serie $\sum_{k=1}^{\infty} f_k$ converge uniformemente en el intervalo $[a, b]$ y las funciones f_k son continuas en $[a, b]$, entonces

$$\sum_{k=1}^{\infty} \int_a^b f_k(x)\,dx = \int_a^b \left[\sum_{k=1}^{\infty} f_k(x)\right] dx.$$

Demostración. Por el Teorema 8.3 la función $\sum_{k=1}^{\infty} f_k(x)$ es continua y, por tanto, integrable, en $[a, b]$. Tenemos que

$$\int_a^b \left[\sum_{k=1}^{\infty} f_k(x)\right] dx - \sum_{k=1}^{N} \int_a^b f_k(x)\,dx = \int_a^b \sum_{k=N+1}^{\infty} f_k(x)\,dx. \tag{2}$$

Sea $\epsilon > 0$ y elijamos K de modo que si $N > K$, entonces $\left|\sum_{k=N+1}^{\infty} f_k(x)\right| < \epsilon(b-a)^{-1}$, para todo x en $[a, b]$. Entonces como para toda función continua g se tiene que

$$\left| \int_a^b g(x)\,dx \right| \le (b-a) \max_{a\le x\le b} |g(x)|,$$

tenemos que

$$\left| \int_a^b \sum_{k=N+1}^{\infty} f_k(x)\,dx \right| \le (b-a)\cdot\epsilon\cdot(b-a)^{-1} = \epsilon, \quad \text{para} \quad N > K.$$

Por tanto, el valor absoluto del lado izquierdo de la Ec. (2) es menor que ϵ para $N > K$, como se quería demostrar.

El intercambio de la derivación con la suma de una serie requiere un número mayor de hipótesis que los teoremas anteriores.

8.5 Teorema

Sea $f_1, f_2, f_3, \ldots$ una sucesión de funciones continuamente diferenciables definidas en un intervalo $[a, b]$. Si $\sum_{k=1}^{\infty} f_k(x) = f(x)$ para todo x en $[a, b]$ (convergencia puntual) y $\sum_{k=1}^{\infty} df_k/dx$ converge uniformemente en $[a, b]$, entonces f es continuamente diferenciable y

$$\frac{d}{dx} \sum_{k=1}^{\infty} f_k(x) = \sum_{k=1}^{\infty} \frac{df_k}{dx}(x).$$

Demostración. Por el teorema fundamental del cálculo,

$$\sum_{k=1}^{N} [f_k(x) - f_k(a)] = \sum_{k=1}^{N} \int_a^x f_k'(t)\,dt$$

$$= \int_a^x \left[\sum_{k=1}^{N} f_k'(t)\right] dt. \tag{3}$$

Haciendo tender N a infinito, tenemos que $\sum_{k=1}^{\infty} f_k(x) = f(x)$; de modo que

$$f(x) - f(a) = \int_a^x \left[\sum_{k=1}^{\infty} f_k'(t)\right] dt,$$

donde hemos usado la convergencia puntual en el lado izquierdo de la Ec. (3) y, en el lado derecho, la convergencia uniforme y el Teorema 8.4. Derivando ambos lados de la última ecuación se tiene que

$$f'(x) = \sum_{k=1}^{\infty} f_k'(x),$$

que es la conclusión del teorema.

Ejemplo 3. Consideremos la serie trigonométrica

$$\sum_{k=1}^{\infty} \frac{\operatorname{sen} kx}{k^4}.$$

Es claro que esta serie converge para todo x real. Además la serie de las derivadas de los términos de la serie original es

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^3}.$$

Usando el criterio de Weierstrass se ve que esta serie converge uniformemente puesto que

$$\left|\frac{\cos kx}{k^3}\right| \leq \frac{1}{k^3},$$

y $\sum_1^\infty (1/k^3)$ converge. Por tanto, por el Teorema 8.5,

$$\frac{d}{dx}\sum_{k=1}^{\infty}\frac{\operatorname{sen} kx}{k^4} = \sum_{k=1}^{\infty}\frac{\cos kx}{k^3}.$$

El mismo argumento puede ser obtenido de nuevo, obteniéndose

$$\frac{d^2}{dx^2}\sum_{k=1}^{\infty}\frac{\operatorname{sen} kx}{k^4} = -\sum_{k=1}^{\infty}\frac{\operatorname{sen} kx}{k^2}.$$

EJERCICIOS

1. Demuestre que la serie $\sum_{k=0}^{\infty} x^k$ converge uniformemente para $-d \leq x \leq d$, donde $0 < d < 1$.

2. (a) Demuestre que la serie trigonométrica $\sum_{k=1}^{\infty}\frac{\cos kx}{k^2}$ converge uniformemente para todo x real.
   (b) Demuestre que la serie de la parte (a) define una función continua para todo x real.

3. (a) Demuestre que si una serie trigonométrica

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \operatorname{sen} kx$$

converge uniformemente en $[-\pi, \pi]$, entonces ella converge uniformemente para *todo* x real.
   (b) Demuestre que la serie uniformemente convergente de la parte (a) es la serie de Fourier de la función que ella representa. [*Indicación*: Use el Teorema 8.4.]

4. (a) Demuestre que si $|c_k| \leq B$ para algún número B fijo, entonces la serie

$$u(x, t) = \sum_{k=1}^{\infty} c_k e^{-k^2 t} \operatorname{sen} kx$$

es una solución de la ecuación diferencial

$$u_{xx} = u_t \quad \text{para} \quad t > 0 \quad \text{y } x \text{ en } [0, \pi],$$

que satisface $u(0, t) = u(\pi, t) = 0$. [*Indicación*: Use el Teorema 8.5.]
   (b) Demuestre que si $u(x, t)$ de la parte (a) está definido para $t = 0$ por una serie uniformemente convergente, entonces $u(x, t)$ es continua en el subconjunto S de $\mathcal{R}^2$ definido por $0 \leq t$, $0 \leq x \leq \pi$.
   (c) Demuestre que la función $u(x, t)$ es infinitamente diferenciable con respecto de x y de t, para $t > 0$.

5. Demuestre que si una serie trigonométrica de la forma indicada en el Ejercicio 3(a) satisface la condición $|a_n| \leq A/n^2$, $|b_n| \leq B/n^2$ para $n = 1, 2, 3, \ldots$, donde A y B son constantes, entonces la serie trigonométrica es una serie de Fourier.

6. Considerando las sumas parciales de la serie de potencias $\sum_{k=0}^{\infty} x^k$ para $-1 < x < 1$, demuestre que la serie no converge uniformemente en $(-1, 1)$.

7. Demuestre que $\sum_{k=1}^{\infty}(-1)^k(1-x)x^k$ converge uniformemente en $[0, 1]$ y que $\sum_{k=1}^{\infty}(1-x)x^k$ solo converge puntualmente en $[0, 1]$.

8. (a) Suponga que las series $\sum_{k=1}^{\infty} k^2 a_n$ y $\sum_{k=1}^{\infty} k^2 b_n$ convergen absolutamente. Demuestre que

$$w(x, t) = \sum_{k=1}^{\infty} \operatorname{sen} kx(a_k \cos kat + b_k \operatorname{sen} kat)$$

es una solución de la ecuación de ondas en una dimensión $a^2 w_{xx} = w_{tt}$. [*Indicación*: Use el criterio de Weierstrass y el Teorema 8.5.]
   (b) Demuestre que la solución $w(x, t)$ de la parte (a) satisface las condiciones de frontera $w(0, t) = w(\pi, t) = 0$ para $t \geq 0$ y la condición inicial $w(x, 0) = h(x)$, donde h es una función dos veces continuamente diferenciable.

9. Demuestre que si la convergencia uniforme se remplaza por convergencia puntual en los Teoremas: (a) 8.3, (b) 8.4 y (c) 8.5, las afirmaciones resultantes son falsas.

10. ¿Pueden ser demostrados los Teoremas 8.4 y 8.5 si suponemos que las funciones involucradas tienen valores en $\mathcal{R}^n$? [*Indicación*: Véanse los Ejercicios 16 y 17 de la Sec. 2 del Cap. 3.]

9. Funciones ortogonales

Algunas de las propiedades del desarrollo de Fourier son compartidas por una clase bastante grande de desarrollos similares en los cuales las sucesiones $\{\operatorname{sen} kx\}$ y $\{\cos kx\}$ son remplazadas por alguna sucesión $\{\varphi_k\}_{k=1,2,\ldots}$ de funciones mutuamente ortogonales. La ortogonalidad es considerada con respecto de algún producto interior en un espacio vectorial de funciones. Por ejemplo, si f y g son elementos del espacio $C[-\pi, \pi]$ de las funciones continuas en el intervalo $[-\pi, \pi]$, podemos definir su producto interno por

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx. \tag{1}$$

Se puede comprobar fácilmente (comparar con el Teorema 5.3) que si definimos $\varphi_1(x) = 1/\sqrt{2\pi}$, $\varphi_{2n}(x) = (\cos nx)/\sqrt{\pi}$ y $\varphi_{2n+1}(x) = (\operatorname{sen} nx)/\sqrt{\pi}$, entonces. La sucesión $\{\varphi_k\}_{k=1,2,\ldots}$ es **ortonormal** con respecto del producto interno

$$\langle \varphi_k, \varphi_l\rangle = \begin{cases} 1, & k = l \\ 0, & k \neq l. \end{cases}$$

definido por la Ec. (1). El término «normal» se usa ya que las funciones han sido normalizadas requiriendo que $\|\varphi_k\| = \langle \varphi_k, \varphi_k \rangle^{1/2} = 1$. En el caso trigonométrico la normalización se logra dividiendo por $\sqrt{\pi}$.

Para ver la importancia de las sucesiones ortonormales generales, consideramos el siguiente problema: Sea $\langle f, g \rangle$ un producto interno en un espacio vectorial y $\{\varphi_k\}_{k=1,2,\ldots}$ una sucesión de elementos ortonormales con respecto del producto interno. Usando la norma definida por $\|f\| = \langle f, f \rangle^{1/2}$, tratamos de determinar coeficientes c_k, $k = 1, 2, \ldots, N$, tales que

$$\left\| g - \sum_{k=1}^{N} c_k \varphi_k \right\|$$

sea mínimo para g y N dados. Como la sucesión φ_k es ortonormal, la solución es muy sencilla. En efecto, sumando y sustrayendo $\sum_{k=1}^{N} \langle g, \varphi_k \rangle^2$, tenemos que

$$\begin{aligned} 0 \leq \left\| g - \sum_{k=1}^{N} c_k \varphi_k \right\|^2 &= \left\langle g - \sum_{k=1}^{N} c_k \varphi_k,\ g - \sum_{k=1}^{N} c_k \varphi_k \right\rangle \\ &= \|g\|^2 - 2 \sum_{k=1}^{N} c_k \langle g, \varphi_k \rangle + \sum_{k=1}^{N} c_k^2 \\ &= \|g\| \quad - \sum_{k=1}^{N} \langle g, \varphi_k \rangle^2 \\ &\quad + \sum_{k=1}^{N} [\langle g, \varphi_k \rangle^2 - 2 c_k \langle g, \varphi_k \rangle + c_k^2] \\ &= \|g\|^2 - \sum_{k=1}^{N} \langle g, \varphi_k \rangle^2 + \sum_{k=1}^{N} [\langle g, \varphi_k \rangle - c_k]^2. \end{aligned} \tag{2}$$

Pero los dos primeros términos de la última expresión son independientes de la elección de los c_k y la última suma queda minimizada si se toma $c_k = \langle g, \varphi_k \rangle$. Los números $\langle g, \varphi_k \rangle$ se llaman los **coeficientes de Fourier** de g con respecto de la sucesión ortonormal $\{\varphi_k\}$. Nótese que la simplicidad de la respuesta depende totalmente de la ortogonalidad de $\{\varphi_k\}$.

Resumimos lo que hemos demostrado como sigue:

9.1 Teorema

Sea $\{\varphi_k\}_{k=1,2,\ldots}$ una sucesión ortonormal en un espacio vectorial con un producto interno. Entonces, dado un elemento g del espacio, la distancia

$$\left\| g - \sum_{k=1}^{N} c_k \varphi_k \right\|$$

es mínima para $N = 1, 2, \ldots$ si se toman los c_k como los coeficientes de Fourier $\langle g, \varphi_k \rangle$.

La importancia de la conclusión del Teorema 9.1 es que los c_k están determinados unívocamente, independientemente de N. En otras palabras, si queremos mejorar la aproximación aumentando N, entonces el Teorema 9.1 dice que los c_k ya calculados no cambian y que solo es necesario calcular los coeficientes adicionales $c_{N+1} = \langle g, \varphi_{N+1} \rangle$, etc.

Como subproducto de la demostración del Teorema 9.1, tenemos

9.2 Desigualdad de Bessel

$$\|g\|^2 \geq \sum_{k=1}^{\infty} \langle g, \varphi_k \rangle^2.$$

Demostración. La desigualdad

$$0 \leq \|g\|^2 - \sum_{k=1}^{N} \langle g, \varphi_k \rangle^2 + \sum_{k=1}^{N} [\langle g, \varphi_k \rangle - c_k]^2$$

ha sido establecida en el último paso de la Ec. (2). Si tomamos $c_k = \langle g, \varphi_k \rangle$, esta desigualdad se transforma en

$$0 \leq \|g\|^2 - \sum_{k=1}^{N} \langle g, \varphi_k \rangle^2$$

La desigualdad de Bessel se obtiene haciendo tender N a infinito.

Ejemplo 1. La aproximación de g por una suma $\sum_{k=1}^{N} c_k \varphi_k$ ha sido medida por su norma en el Teorema 9.1. Para ver el significado de esto en el caso de polinomios trigonométricos, usamos el producto interno de la Ec. (1) en el espacio $C[-\pi, \pi]$ de las funciones continuas en $[-\pi, \pi]$. Dada la sucesión ortonormal $\varphi_1(x) = 1/\sqrt{2\pi}$, $\varphi_{2n}(x) = (\cos nx)/\sqrt{\pi}$, $\varphi_{2n+1}(x) = (\operatorname{sen} nx)/\sqrt{\pi}$, entonces para g en $C[-\pi, \pi]$, tratamos de minimizar la norma

$$\left\| g - \sum_{k=1}^{2N+1} c_k \varphi_k \right\|.$$

Hemos visto que esto se hace tomando $c_k = \langle g, \varphi_k \rangle$. Pero, por definición del producto interno,

$$\langle g, \varphi_k \rangle = \begin{cases} \displaystyle\int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} g(t)\, dt, & k = 1 \\ \displaystyle\int_{-\pi}^{\pi} g(t) \frac{\cos nt}{\sqrt{\pi}}\, dt, & k = 2n \\ \displaystyle\int_{-\pi}^{\pi} g(t) \frac{\operatorname{sen} nt}{\sqrt{\pi}}\, dt, & k = 2n + 1. \end{cases}$$

Por tanto, los términos $c_k\varphi_k$ son

$$\langle g, \varphi_1\rangle\varphi_1(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\,dt,$$

$$\langle g, \varphi_{2n}\rangle\varphi_{2n}(x) = \left[\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos nt\,dt\right]\cos nx,$$

$$\langle g, \varphi_{2n+1}\rangle\varphi_{2n+1}(x) = \left[\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\operatorname{sen} nt\,dt\right]\operatorname{sen} nx.$$

Entonces

$$\sum_{k=1}^{2N+1} c_k\varphi_k(x) = \frac{a_0}{2} + \sum_{k=1}^{N}(a_k\cos kx + b_k\operatorname{sen} kx),$$

donde a_k y b_k son los coeficientes de Fourier trigonométricos como se definieron en la Sec. 5. El cuadrado de la norma que se desea minimizar tiene la forma

$$\int_{-\pi}^{\pi}\left[g(x) - \frac{a_0}{2} - \sum_{k=1}^{N}(a_k\cos kx + b_k\operatorname{sen} kx)\right]^2 dx,$$

y el Teorema 9.1 dice que, por N fijo, este mínimo se consigue tomando para a_k y b_k los coeficientes de Fourier de g.

La minimización de una integral de la forma

$$\int_a^b\left[g(x) - \sum_{k=1}^{N} c_k\varphi_k(x)\right]^2 dx$$

se llama la mejor **aproximación en promedio cuadrado** de g. En este sentido podemos decir que la aproximación de Fourier es la mejor aproximación en promedio cuadrado por un polinomio trigonométrico.

En la Sec. 5 hicimos notar que las funciones cos kx y sen kx son soluciones de la ecuación diferencial $y'' = -k^2y$. Este resultado puede expresarse diciendo que cos kx y sen kx son **funciones propias** del operador diferencial $d^2/(dx^2)$ que corresponde al **valor propio** $-k^2$. Más generalmente, sea L una transformación lineal definida en un espacio vectorial con valores en el mismo espacio. Decimos que un vector f *distinto de cero* es un **vector propio** de L correspondiente al **valor propio** λ si $Lf = \lambda f$, donde λ es un número real.

Para ver la relación entre vectores propios y conjuntos ortogonales de funciones, necesitamos otra definición. Sea $\mathfrak{F}$ un espacio vectorial con producto interno $\langle f, g\rangle$. Sea L una transformación lineal de $\mathfrak{F}$ en $\mathfrak{F}$. Entonces L es **simétrica** con respecto del producto interno si $\langle Lf, g\rangle = \langle f, Lg\rangle$ para todos los vectores f y g tales que Lf y Lg estén definidos como elementos de $\mathfrak{F}$. Una transformación lineal con espacio dominio y espacio imagen iguales se llama un **operador lineal**.

Ejemplo 2. Sea $\langle f, g\rangle$ el producto interior en $C[-\pi, \pi]$ definido por la Ec. (1). Si definimos $Lf = d^2f/dx^2$, entonces es claro que Lf está en $C[-\pi, \pi]$

solo si f tiene segunda derivada continua. Por tanto, para que L sea simétrico es necesario que $\langle Lf, g\rangle = \langle f, Lg\rangle$ para todas las funciones f y g con segunda derivada continua en $[-\pi, \pi]$. En otras palabras, se debe tener que

$$\int_{-\pi}^{\pi} f''(x)g(x)\,dx = \int_{-\pi}^{\pi} f(x)g''(x)\,dx.$$

Integrando dos veces por partes, se tiene que

$$\int_{-\pi}^{\pi} f''(x)g(x)\,dx = f'(\pi)g(\pi) - f'(-\pi)g(-\pi) - f(\pi)g'(\pi) + f(-\pi)g'(-\pi) + \int_{-\pi}^{\pi} f(x)g''(x)\,dx.$$

Por tanto, para hacer L simétrico restringimos su dominio a un subespacio de $C[-\pi, \pi]$ para cuyas funciones los términos no integrados tienen suma cero. Esto puede hacerse de varias maneras. Por ejemplo, podemos restringir L al subespacio formado por las funciones h en $C[-\pi, \pi]$ tales que $h(\pi) = h(-\pi) = 0$ o al subespacio de aquellos que satisfacen $h'(\pi) = h'(-\pi) = 0$. Con cualquiera de estas restricciones L es simétrico. Nótese que las restricciones del tipo requerido son condiciones de frontera puesto que especifican los valores de h en los extremos π y $-\pi$, del dominio de definición.

La conexión entre funciones ortogonales y operadores simétricos es la siguiente:

9.3 Teorema

Sea L un operador lineal simétrico definido en un espacio vectorial $\mathfrak{F}$ con un producto interno. Si f_1 y f_2 son vectores propios de L que corresponden a valores propios distintos λ_1 y λ_2, entonces f_1 y f_2 son ortogonales.

Demostración. Supongamos que

$$Lf_1 = \lambda_1 f_1, \qquad Lf_2 = \lambda_2 f_2,$$

y demostremos que $\langle f_1, f_2\rangle = 0$. Tenemos que

$$\langle Lf_1, f_2\rangle = \langle \lambda_1 f_1, f_2\rangle = \lambda_1\langle f_1, f_2\rangle$$

y

$$\langle f_1, Lf_2\rangle = \langle f_1, \lambda_2 f_2\rangle = \lambda_2\langle f_1, f_2\rangle.$$

Como L es simétrico, $\langle Lf_1, f_2\rangle = \langle f_1, Lf_2\rangle$ de modo que $\lambda_1\langle f_1, f_2\rangle = \lambda_2\langle f_1, f_2\rangle$ o $(\lambda_1 - \lambda_2)\langle f_1, f_2\rangle = 0$. Como λ_1 y λ_2 son distintos, debemos tener que $\langle f_1, f_2\rangle = 0$.

Consideramos el operador diferencial de **Sturm-Liouville**:

$$Lf = (pf') + qf, \tag{3}$$

donde p es una función continuamente diferenciable y q una función continua. (El operador d^2/dx^2 es un caso especial en que $p(x) \equiv 1$ y $q(x) \equiv 0$.) Queremos ver qué condiciones de frontera deben imponerse al dominio de L para que L sea simétrico con respecto del producto interno,

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx.$$

La siguiente fórmula simplifica el problema.

9.4 Fórmula de Lagrange

Si L está dado por $Lf = (pf')' + qf$ en un intervalo $[a, b]$, entonces

$$\langle f_1, Lf_2\rangle - \langle Lf_1, f_2\rangle = [p(f_1f_2' - f_1'f_2)]_a^b.$$

Demostración. Partiendo de la definición de L, reordenamos $f_1(Lf_2) - (Lf_1)f_2$ como sigue:

$$\begin{aligned} f_1(Lf_2) - (Lf_1)f_2 &= f_1[(pf_2')' + qf_2] - f_2[(pf_1')' + qf_1] \\ &= f_1(pf_2')' - f_2(pf_1')' \\ &= f_1[pf_2'' + p'f_2'] - f_2[pf_1'' + p'f_1'] \\ &= p'[f_1f_2' - f_1'f_2] + p[f_1f_2'' - f_1''f_2] \\ &= [p(f_1f_2' - f_1'f_2)]'. \end{aligned}$$

Integrando ambos lados entre a y b se tiene la fórmula de Lagrange.

La Ec. 9.4 muestra que toda condición en el coeficiente p o en el espacio que contiene a f_1 y f_2 que asegure que $[p(f_1f_2' - f_1'f_2)]_a^b = 0$ hará simétrico al operador L.

Ejemplo 3. El operador L definido por $Lf(x) = (1 - x^2)f''(x) - 2xf'(x)$ es de la forma de la Ec. (3), donde si se toma $p(x) = (1 - x^2)$ y $q(x) \equiv 0$. Consideraremos que L opera en las funciones con segunda derivada continua en $[-1, 1]$. Para hacer L simétrico debemos asegurarnos que el lado derecho de la fórmula de Lagrange es cero para $a = -1$, $b = 1$. Pero $p(x) = (1 - x^2)$, de modo que $p(-1) = p(1) = 0$. Por tanto, L es simétrico en el espacio $C[-1, 1]$ y su dominio está formado por las funciones en $C[-1, 1]$ con segunda derivada continua.

El operador simétrico Lf definido en el ejemplo está asociado con la ecuación diferencial

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0. \tag{4}$$

que se llama la **ecuación de Legendre** de índice n y es satisfecha por el n-ésimo **polinomio de Legendre** definido por

$$\mathbf{9.5} \qquad P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n, \qquad n = 0, 1, 2, \ldots.$$

Derivando sucesivamente, se comprueba que P^n satisface la Ec. (4). Véase el Ejercicio 8. La importancia del hecho de que P^n satisface la ecuación de Legendre se ve si escribimos la ecuación en la forma $Ly = -n(n + 1)y$, donde L es el operador simétrico $Ly = (1 - x^2)y'' - 2xy'$ en $C[-1, 1]$. Entonces P_n puede ser considerado como una función propia de L que corresponde al valor propio $-n(n + 1)$. Por tanto, por el Teorema 9.3, los polinomios de Legendre son ortogonales. Es decir,

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = 0, \qquad n \neq m.$$

Además, un cálculo más o menos complicado (véase Ejercicio 10), muestra que

$$\int_{-1}^{1} P_n^2(x)\,dx = \frac{2}{2n + 1}, \qquad n = 0, 1, 2, \ldots.$$

Por tanto, la sucesión normalizada $\{(\sqrt{2n + 1/2})P^n(x)\}$ $n = 0, 1, 2, \ldots$ es una sucesión ortonormal en $C[-1, 1]$.

La **aproximación N-ésima de Fourier-Legendre** de una función g en $C[-1, 1]$ es la suma finita

$$\sum_{k=0}^{N} c_kP_k(x),$$

donde

$$c_k = \frac{2k + 1}{2}\int_{-1}^{1} g(x)P_k(x)\,dx. \tag{5}$$

Entonces el Teorema 9.1 implica que la mejor aproximación promedio cuadrado de g por una combinación de polinomios de Legendre está dada por la aproximación de Fourier-Legendre. En otras palabras,

$$\int_{-1}^{1}\left[g(x) - \sum_{k=0}^{N} c_kP_k(x)\right]^2 dx$$

es minimizada por los c_k de la Ec. (5).

EJERCICIOS

1. (a) Verifique que

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$$

define un producto interno en el espacio $C[a, b]$ de las funciones continuas definidas en $[a, b]$.

(b) ¿Qué condición debe satisfacer una función w definida en $[a, b]$ para que

$$\langle f, g\rangle = \int_a^b f(x)g(x)w(x)\,dx$$

defina un producto interno?

2. Sea $\{\varphi_k\}$ una sucesión ortonormal de funciones en $C[a, b]$ y sea g una función en $C[a, b]$. Demuestre que si la función real

$$\Delta(c_1, \ldots, c_N) = \int_a^b \left[g(x) - \sum_{k=1}^{N} c_k \varphi_k(x) \right] dx$$

tiene un mínimo local como función de $(c_1, \ldots, c_N)$, entonces

$$c_k = \int_a^b g(x)\varphi_k(x)\, dx.$$

[*Indicación*: Derive bajo el signo integral la fórmula para Δ. Esto queda justificado por las hipótesis consideradas. (Véase el Ejercicio 7 de la Sec. 2 del Cap. 6).]

3. Sea $\{\varphi_k\}$ una sucesión ortonormal de funciones en $C[a, b]$ y sea g en $C[a, b]$. Use la desigualdad de Bessel para demostrar

(a) $\sum_{k=1}^{\infty} \left[\int_a^b g(x)\varphi_k(x)\, dx \right]^2$ converja.

(b) Si c_k es el coeficiente k-ésimo de Fourier de g con respecto de $\{\varphi_k\}$ entonces $\lim_{k\to\infty} c_k = 0$.

(c) Encuentre un ejemplo de una función trigonométrica cuya serie trigonométrica de Fourier tenga coeficiente a_k y b_k tales que $\sum_{k=1}^{\infty} (a_k^2 + b_k^2)$ converja sin que $\sum_{k=1}^{\infty} a_k$ o $\sum_{k=2}^{\infty} b_k$ converjan.

4. Sea L una transformación lineal de un espacio vectorial $\mathfrak{F}_1$ en otro espacio $\mathfrak{F}_2$ que no tiene relación con $\mathfrak{F}_1$. ¿Puede L tener vectores propios?

5. Encuentre todas las funciones propias y sus correspondientes valores propios del operador diferencial d^2/dx^2 que satisfacen cada una de las condiciones de frontera siguientes. Es decir, resuelva $y'' = \lambda y$ con las condiciones de frontera

(a) $y(0) = y(\pi) = 0$.
(b) $y(0) = y(\pi) + y'(\pi) = 0$.
(c) $y(0) = y'(\pi) = 0$.

6. Sea $C[a, b]$ el conjunto de las funciones continuas definidas en $[a, b]$, $C'[a, b]$ el conjunto de las funciones continuamente diferenciables en $[a, b]$ y $C''[a, b]$ el de las funciones con segunda derivada continua.

(a) Demuestre que $C[a, b]$, $C'[a, b]$ y $C''[a, b]$ son espacios vectoriales, cada uno de ellos contenido en el siguiente.

(b) Sea $B_{x_0}[a, b]$ el conjunto de las funciones f en $C[a, b]$ que satisfacen una condición de la forma

$$cf(x_0) + df'(x_0) = 0,$$

donde c y d son constantes tales que ambos no son iguales a cero. Demuestre que $B_{x_0}[a, b]$ es un subespacio vectorial de $C[a, b]$.

(c) Demuestre que $C''[a, b] \cap B_a[a, b] \cap B_b[a, b]$ es un subespacio vectorial de $C[a, b]$.

7. Demuestre que el operador diferencial d^2/dx^2 es simétrico con respecto del producto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\, dx,$$

y con respecto de los siguientes conjuntos de condiciones de frontera

(a) $f(-1) = f(1) = 0$.
(b) $f'(1) = f'(-1) = 0$.
(c) $c_1 f(-1) + d_1 f'(-1) = c_2 f(1) + d_2 f'(1) = 0$,
donde $c_i^2 + d_i^2 > 0$.

8. Verifique que el polinomio de Legendre P_n definido por la Ec. 9.5 satisface la ecuación de Legendre: $(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$. [*Indicación*: Sea $u = (x^2 - 1)^n$. Entonces $(x^2 - 1)u' = 2nxu$. Derive esta ecuación $(n + 1)$ veces con respecto de x.]

9. Calcule los polinomios de Legendre P_0, P_1 y P_2.

10. (a) Use la Ec. 9.5 y varias integraciones por partes para demostrar que

$$\int_{-1}^{1} P_n^2(x)\, dx = \frac{(2n)!}{2^{2n}(n!)^2} \int_{-1}^{1} (1 - x^2)^n\, dx.$$

(b) Demuestre que

$$\int_{-1}^{1} (1 - x^2)^n\, dx = 2\int_0^{\pi/2} \operatorname{sen}^{2n+1}\theta\, d\theta.$$

(c) Demuestre que

$$\int_0^{\pi/2} \operatorname{sen}^{2n+1}\theta\, d\theta = \frac{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n)}{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n + 1)}.$$

(d) Demuestre que

$$\int_{-1}^{1} P_n^2(x)\, dx = \frac{2}{2n + 1}.$$

11. Demuestre que el n-ésimo polinomio de Legendre, P_n, tiene n raíces distintas en el intervalo $[-1, 1]$. [*Indicación*: Use la Ec. 9.5 y el teorema de Rolle.]

12. La ecuación de Laplace en tres dimensiones tiene la forma

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial u}{\partial r}\right) + \frac{1}{\operatorname{sen}\varphi} \frac{\partial}{d\varphi}\left(\operatorname{sen}\varphi \frac{\partial u}{\partial \varphi}\right) + \frac{1}{\operatorname{sen}^2\varphi} \frac{\partial^2 u}{\partial \theta^2} = 0.$$

en coordenadas polares (r, φ, θ).

(a) Demuestre que para soluciones $u(r, \varphi, \theta) = v(r, \varphi)$ que son independientes de θ, la ecuación tiene la forma

$$r \frac{\partial^2}{\partial r^2}(rv) + \frac{1}{\operatorname{sen}\varphi} \frac{\partial}{\partial \varphi}\left(\operatorname{sen}\varphi \frac{\partial v}{\partial \varphi}\right) = 0.$$

(b) Demuestre que el método de la separación de variables aplicada a la ecuación de la parte (a) conduce a las ecuaciones diferenciales ordinarias

$$r^2 G'' + 2rG' = \lambda G,$$

$$\frac{1}{\operatorname{sen}\varphi} \frac{d}{d\varphi}\left(\operatorname{sen}\varphi \frac{dH}{d\varphi}\right) = -\lambda H.$$

(c) Demuestre que la ecuación para $G(r)$ tiene soluciones

$$r^{-(1/2)+\sqrt{\lambda+(1/4)}} \quad \text{y} \quad r^{-(1/2)-\sqrt{\lambda+(1/4)}}. \qquad [\textit{Sugerencia}: \text{Sea } r = e^t.]$$

(d) Demuestre que la ecuación para H puede ser transformada en una ecuación de Legendre

$$(1 - x^2)H'' - 2xH' + \lambda H = 0.$$

[*Indicación*: Tome $x = \cos \varphi$.]

(e) Tomando $\lambda = n(n + 1)$ encuentre una sucesión de soluciones de la ecuación diferencial de la parte (a).

13. Demuestre que para la sucesión ortonormal deducida de los polinomios de Legendre, el desarrollo general $\sum_{k=0}^{N} \langle g, \varphi_k \rangle \varphi_k$ se reduce a $\sum_{k=0}^{N} c_k P_n(x)$, donde c_k está dado por la Ec. (5).

14. Use el Teorema 8.2 del Cap. 2 para demostrar que los polinomios de Legendre normalizados

$$\sqrt{\frac{2n+1}{2}} P_n(x)$$

son iguales a la sucesión de polinomios obtenidos aplicando el método de Gram-Schmidt a las funciones $\{1, x, x^2, \ldots\}$. Use el producto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\, dx$$

definido en $C[0, 1]$.

6

Integrales múltiples

1. Integrales iteradas

Este capítulo está dedicado al estudio de la integral de funciones reales con dominio en $\mathcal{R}^n$. Integrales de este tipo aparecen en varias ramas de la matemática pura y aplicada con interpretaciones tales como volumen, masa, probabilidad y flujo. En esta sección empezamos con las integrales iteradas, ya que ellas son útiles para las computaciones y constituyen una transición de la integral definida ordinaria

$$\int_a^b f(x)\, dx,$$

de una función real de una variable real a la integral múltiple. Empezamos con $n = 2$, es decir, con la integral iterada de funciones $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$.

Supongamos que $f(x, y)$ es una función definida en un rectángulo $a \leq x \leq b$, $c \leq y \leq d$. Por

$$\int_a^d f(x, y)\, dt$$

entenderemos simplemente la integral de la función de una variable obtenida manteniendo x fijo; por ejemplo,

$$\int_0^2 x^3 y^2\, dy = \left[\frac{x^3 y^3}{3}\right]_{y=0}^{y=2} = \frac{8}{3} x^3.$$

Como lo muestra este ejemplo, si la integral existe, ella depende de x. Por tanto, podemos hacer

$$F(x) = \int_c^d f(x, y)\, dy$$

y definir la **integral iterada**

$$\int_a^b F(x)\, dx = \int_a^b \left[\int_c^d f(x, y)\, dy \right] dx.$$

Una convención usual de notación que adoptaremos es la de omitir los paréntesis y escribir la integral iterada como

$$\int_a^b dx \int_c^d f(x, y)\, dy.$$

Esta notación tiene la conveniencia de indicar qué variable se asocia a cada signo integral: x con $\int_a^b$, y con $\int_c^d$.

Ejemplo 1. Consideremos $f(x, y) = x^2 + y$, definida en la región rectangular $0 \leq x \leq 1$, $1 \leq y \leq 2$.

$$\int_0^1 dx \int_1^2 (x^2 + y)\, dy = \int_0^1 \left[x^2 y + \frac{y^2}{2} \right]_{y=1}^{y=2} dx$$

$$= \int_0^1 [(2x^2 + 2) - (x^2 + \tfrac{1}{2})]\, dx$$

$$= \int_0^1 (x^2 + \tfrac{3}{2})\, dx = \tfrac{1}{3} + \tfrac{3}{2} = \tfrac{11}{6}.$$

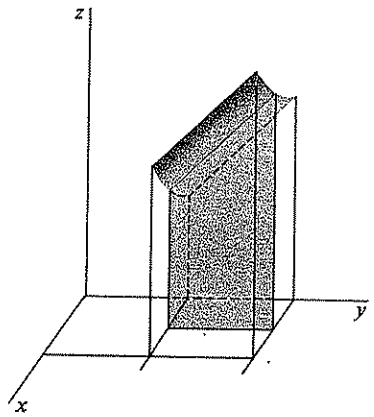


Figura 1

Para interpretar geométricamente este ejemplo, consideraremos la superficie definida por $z = x^2 + y$, mostrada en la Fig. 1. Para cada x en el intervalo entre 0 y 1, la integral

$$\int_1^2 (x^2 + y)\, dy = x^2 + \tfrac{3}{2}$$

es el área de la sección sombreada. Es habitual interpretar como un volumen la integral definida de una función cuyos valores son áreas. Así, es natural considerar la integral iterada

$$\int_0^1 dx \int_1^2 (x^2 + y)\, dy = \tfrac{11}{6}$$

como el volumen de la región tridimensional entre la superficie y el rectángulo $0 \leq x \leq 1$, $1 \leq y \leq 2$.

Ejemplo 2. Podemos hacer las integraciones del Ejemplo 1 en el orden inverso

$$\int_1^2 dy \int_0^1 (x^2 + y)\, dx = \int_1^2 \left[\frac{x^3}{3} + yx \right]_{x=0}^{x=1} dy$$

$$= \int_1^2 (\tfrac{1}{3} + y)\, dy = \left[\frac{y}{3} + \frac{y^2}{2} \right]_1^2$$

$$= (\tfrac{2}{3} + 2) - (\tfrac{1}{3} + \tfrac{1}{2}) = \tfrac{11}{6}.$$

Ahora,

$$\int_0^1 (x^2 + y)\, dx = \tfrac{1}{3} + y$$

es el área de la sección paralela al plano xz. Véase Fig. 2. La segunda integral es, nuevamente, el volumen de la región tridimensional entre la superficie $z =$

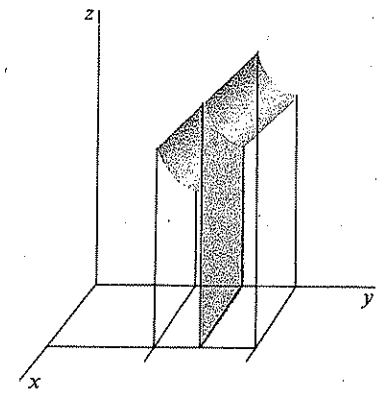


Figura 2

$x^2 + y$ y el rectángulo $0 \le x \le 1$, $1 \le y \le 2$. Por tanto, no es sorprendente que las integrales iteradas de los Ejemplos 1 y 2 sean iguales.

Es importante poder integrar sobre subconjuntos del plano más generales que los rectángulos. En ese tipo de problemas, los límites de la primera integral dependerán de la otra variable.

Ejemplo 3. Consideremos la integral iterada

$$\int_0^1 dx \int_0^{1-x^2} (x + y)\, dy = \int_0^1 \left[xy + \frac{y^2}{2} \right]_0^{1-x^2} dx$$
$$= \int_0^1 \left[x(1 - x^2) + \frac{(1 - x^2)^2}{2} \right] dx$$
$$= \int_0^1 \left[x - x^3 + \frac{1 - 2x^2 + x^4}{2} \right] dx = \tfrac{31}{60}.$$

Para cada x entre 0 y 1, el número y está entre $y = 0$ e $y = 1 - x^2$. En otras palabras, el punto (x, y) varía a lo largo del segmento de recta que une a $(x, 0)$ y $(x, 1 - x^2)$. Al variar x entre 0 y 1, este segmento describe el área sombreada B de la Fig. 3. El integrando $f(x, y) = x + y$ tiene un gráfico (véase Fig. 4) y la integral iterada es el volumen entre este gráfico y B.

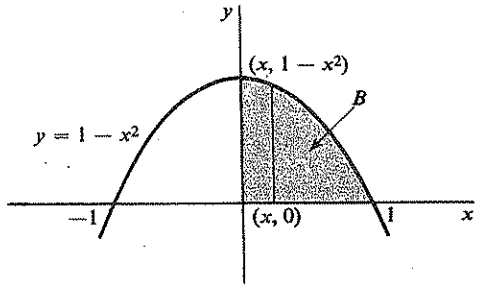


Figura 3

Supongamos dada una integral iterada sobre una región plana B, en la cual el integrando es la función constante $f(x, y) = 1$, para todo (x, y) en B. Entonces la integral puede ser interpretada ya sea como el volumen de una placa de base B y género unitario o, simplemente, como el área de B. Por ejemplo,

$$\int_0^1 dx \int_0^{1-x^2} dy = \tfrac{2}{3}$$

es el área de la región B de la Fig. 3.

Ejemplo 4. Sea f definida por $f(x, y) = x^2y + yx^2$ sobre la región acotada por $y = |x|$, $y = 0$, $x = -1$ y $x = 1$. Véase Fig. 5. Las dos integrales iteradas sobre la región son

$$\int_{-1}^1 dx \int_0^{|x|} (x^2y + xy^2)\, dy$$

y

$$\int_0^1 dy \left[\int_{-1}^{-y} (x^2y + xy^2)\, dx + \int_y^1 (x^2y + xy^2)\, dx \right].$$

La segunda integral es la suma de dos partes, ya que, para y fijo entre 0 y 1, la integral con respecto a x se hace en dos intervalos distintos. El cálculo de estas integrales es sencillo. Obtenemos

$$\int_0^1 \left[\frac{x^3y}{3} + \frac{x^2y^2}{2} \right]_{-1}^{-y} + \left[\frac{x^3y}{3} + \frac{x^2y^2}{2} \right]_y^1 dy = \int_0^1 \tfrac{2}{3}(y - y^4)\, dy = \tfrac{1}{3} - \tfrac{2}{15} = \tfrac{1}{5}.$$

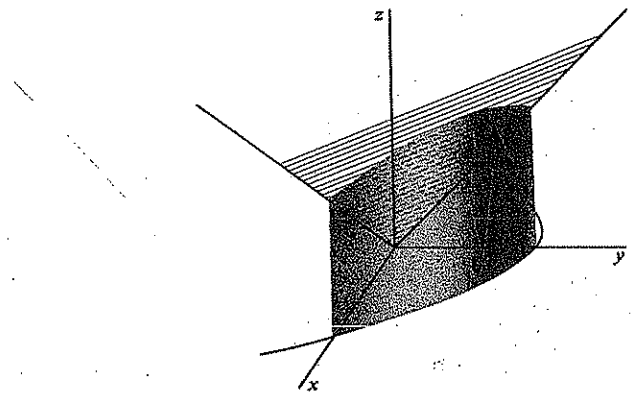


Figura 4

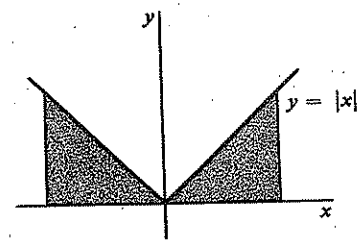


Figura 5

La integral iterada en el otro orden es

$$\int_{-1}^{1}\left[\frac{x^2y^2}{2}+\frac{xy^3}{3}\right]_0^{|x|}dx=\int_{-1}^{1}\left(\frac{x^4}{2}+\frac{x\,|x|^3}{3}\right)dx$$
$$=\int_{-1}^{1}\frac{x^4}{2}\,dx+\int_{-1}^{1}\frac{x\,|x|^3}{3}\,dx.$$

Las funciones $x^4/2$ y $x\,|x|^3/3$ son par e impar, respectivamente. Por tanto,

$$\int_{-1}^{1}\frac{x^4}{2}\,dx+\int_{-1}^{1}\frac{x\,|x|^3}{3}\,dx=\int_0^1 x^4\,dx=\tfrac{1}{5}.$$

En la próxima sección se demostrará el teorema que dice que, bajo condiciones bastante generales, el valor de una integral iterada es independiente del orden de integración. Esto demuestra que el orden de integración en los cálculos de volúmenes no influye en el resultado.

Las integrales iteradas para funciones definidas en subconjuntos de dimensiones mayor que 2 también pueden ser calculadas por repetición de la integración unidimensional.

Ejemplo 5.

$$\int_0^1 dx\int_{x^2}^{x}dy\int_{x}^{2x+y}(x+y+2z)\,dz=\int_0^1 dx\int_{x^2}^{x}(4x^2+6xy+2y)\,dy$$
$$=\int_0^1\left(\tfrac{23}{3}x^3-4x^5-3x-\frac{2x^6}{3}\right)dx$$
$$=\tfrac{23}{12}-\tfrac{4}{6}-\tfrac{1}{2}-\tfrac{2}{21}.$$

No es posible dar una interpretación completa de esta integral mediante dibujos. Sin embargo, la región de integración B puede ser dibujada como en la Fig. 6. Está acotada por arriba por la superficie $z = 2x + y$ y por abajo por $z = x$. En los costados está acotada por las superficies obtenidas proyectando las curvas $y = x^2$ e $y = x$ paralelamente al eje z. Con los mismos límites de integración, la integral

$$\int_0^1 dx\int_{x^2}^{x}dy\int_{x}^{2x+y}dz$$

es el volumen de B. Para x e y fijos, la primera integral

$$\int_{x}^{2x+y}dz,$$

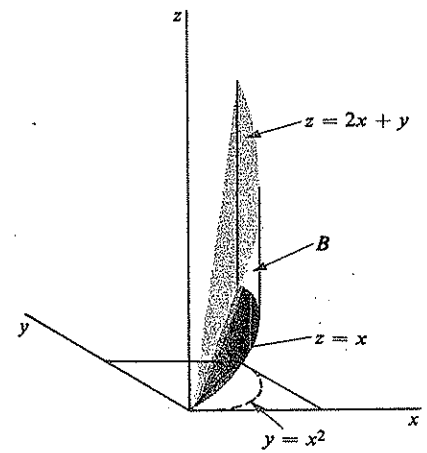


Figura 6

es la longitud del segmento vertical que une los punto (x, y, x) y $(x, y, 2x + y)$. Para x fijo, la integral

$$\int_{x^2}^{x}dy\int_{x}^{2x+y}dz$$

es el área de una sección paralela al plano yz. Finalmente, la integral iterada triple es el volumen.

Ejemplo 6. La integral iterada n veces

$$\int_0^1 dx_1\int_0^{x_1}dx_2\cdots\int_0^{x_{n-1}}dx_n$$

puede ser interpretada como el volumen de la región n-dimensional del espacio euclidiano definida por las desigualdades

$$0\le x_n\le x_{n-1}\le\cdots\le x_2\le x_1\le 1.$$

Para tener una idea de cómo es esta región, consideremos los casos $n = 1$, $n = 2$ y $n = 3$. Para $n = 1$, la integral

$$\int_0^1 dx_1=1$$

es simplemente el largo del intervalo unitario $0 \le x_1 \le 1$. Si $n = 2$, tenemos $0 \le x_2 \le x_1 \le 1$. La región de integración es la intersección de las regiones $0 \le x_2$, $x_2 \le x_1$ y $x_1 \le 1$, mostrados en la Fig. 7. Para $n = 3$, tenemos simultáneamente $0 \le x_3$, $x_3 \le x_2$, $x_2 \le x_1$ y $x_1 \le 1$. Véase Fig. 8. Si denotamos por I_n la integral iterada n veces, entonces $I_1 = 1$, $I_2 = \frac{1}{2}$, $I_3 = \frac{1}{6}$. Estos núme-

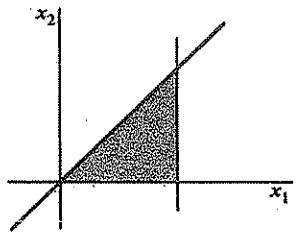


Figura 7

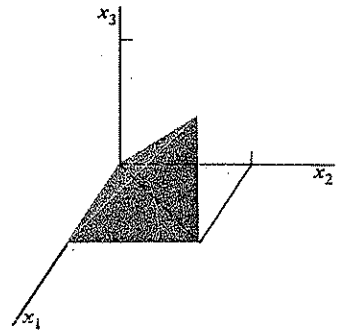


Figura 8

ros pueden ser obtenidos mediante el cálculo de las integrales o notando que ellos son el largo, superficie y volumen, respectivamente, de las regiones de integración. El cálculo de I_n es fácil:

$$\begin{aligned} I_n &= \int_0^1 dx_1 \int_0^{x_1} dx_2 \dots \int_0^{x_{n-1}} dx_n \\ &= \int_0^1 dx_1 \int_0^{x_1} dx_2 \dots \int_0^{x_{n-2}} x_{n-1}\, dx_{n-1} \\ &= \int_0^1 dx_1 \int_0^{x_1} dx_2 \dots \int_0^{x_{n-3}} \frac{x_{n-2}^2}{2}\, dx_{n-2} \\ &= \int_0^1 dx_1 \int_0^{x_1} dx_2 \dots \int_0^{x_{n-4}} \frac{x_{n-3}^3}{3!}\, dx_{n-3} \\ &= \dots = \int_0^1 \frac{x_1^{n-1}}{(n-1)!}\, dx_1 = \frac{1}{n!} \end{aligned}$$

EJERCICIOS

Calcule las siguientes integrales iteradas y dibuje sus regiones de integración.

1. $\int_{-1}^{0} dx \int_{1}^{2} (x^2y^2 + xy^3)\, dy.$
2. $\int_{0}^{2} dy \int_{1}^{3} |x - 2| \operatorname{sen} y\, dx.$ [*Resp*: $1 - \cos 2$.]
3. $\int_{1}^{0} dx \int_{2}^{0} (x + y^2)\, dy.$

4. $\int_{0}^{\pi/2} dy \int_{-y}^{y} \operatorname{sen} x\, dx.$ [*Resp*: 0.]
5. $\int_{-2}^{1} dy \int_{0}^{y^2} (x^2 + y)\, dx.$
6. $\int_{-1}^{1} dx \int_{0}^{|x|} dy.$ [*Resp*: 1.]
7. $\int_{0}^{1} dx \int_{0}^{\sqrt{1-x^2}} dy.$
8. $\int_{1}^{-1} dx \int_{x}^{2x} e^{x+y} dy.$
9. $\int_{0}^{\pi/2} dy \int_{0}^{\cos y} x \operatorname{sen} y\, dx.$
10. $\int_{1}^{2} dx \int_{x^2}^{x^3} x\, dy.$ [*Resp*: $\frac{49}{20}$.]
11. $\int_{0}^{1} dz \int_{0}^{z} dy \int_{0}^{y} dx.$
12. $\int_{0}^{2} dx \int_{1}^{x} dy \int_{2}^{x+y-1} y\, dz.$ [*Resp*: $\frac{2}{3}$.]
13. $\int_{1}^{2} dy \int_{0}^{1} dx \int_{x}^{y} dz.$
14. $\int_{-1}^{1} dx \int_{0}^{|x|} dy \int_{0}^{1} (x + y + z)\, dz.$ [*Resp*: $\frac{5}{6}$.]
15. $\int_{0}^{\pi} \operatorname{sen} x\, dx \int_{0}^{1} dy \int_{0}^{2} (x + y + z)\, dz.$
16. Calcule la integral $\int_{0}^{1} dx \int_{-x}^{x} dy \int_{-x-y}^{x+y} dz \int_{-z}^{x} dw.$
17. Dibuje el subconjunto B definido por $0 \le x \le 1$, $0 \le y \le x$ y escriba y calcule la integral sobre B de la función $f(x, y) = x \operatorname{sen} y$ en los dos órdenes posibles.
18. Dibuje la región definida por $x \ge 0$, $x^2 + y^2 \le 2$ y $x^2 + y^2 \ge 1$. Escriba y calcule la integral sobre esta región de la función $f(x, y) = x^2$ en los dos órdenes posibles. [*Resp*. $3\pi/8$.]
19. Considere dos funciones reales $c(x)$ y $d(x)$ de la variable real x. Suponga que para todo x en el intervalo $a \le x \le b$ tenemos $c(x) \le d(x)$.
    (a) Dibuje dos funciones con estas propiedades y el subconjunto B del plano xy formado por todos los (x, y) tales que $a \le x \le b$ y $c(x) \le y \le d(x)$.
    (b) Exprese el área de B como una integral iterada.
    (c) Escriba la integral iterada de $f(x, y)$ sobre B.
20. Dibuje el subconjunto B de $\mathfrak{R}^3$ definido por $0 \le x \le 1$, $0 \le y \le 1 + x$ y $0 \le z \le 2$.

Escriba y calcule la integral iterada de la función $f(x, y, z) = x^2 + z$ sobre el conjunto B en el orden primero z, luego y y finalmente x. [*Resp.* $\frac{25}{6}$.]

21. Dibuje la región definida por $0 \le x \le 1$, $x^2 \le y \le \sqrt{x}$ y $0 \le z \le x + y$ y calcule la integración iterada, en cualquier orden, de $f(x, y, z) = x + y + z$ sobre esa región.

22. Sea f definida por $f(x, y, z) = 1$ en el hemisferio acotado por el plano $z = 0$ y la superficie $z = \sqrt{1 - x^2 - y^2}$. Calcule una integral iterada de f sobre esta región. [*Resp.* $2\pi/3$.]

23. Sea f definida en el cubo $0 \le x_1 \le 1, 0 \le x_2 \le 1, \ldots, 0 \le x_n \le 1$, por $f(x_1, \ldots, x_n) = x_1 x_2, \ldots, x_n$. Calcule

$$\int_0^1 dx_1 \int_0^1 dx_2 \ldots \int_0^1 x_1 x_2 \ldots x_n \, dx_n.$$

24. Demuestre que si en la integral

$$\int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \ldots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \, dx_n$$

se intercambia el orden de los límites de integración en un número par de signos integrales, entonces el valor de la integral no cambia. Si los límites se intercambian en un número impar de signos integrales, la integral cambia de signo.

25. Calcule

$$\int_0^1 dx_1 \int_0^1 dx_2 \ldots \int_0^1 dx_{n-1} \int_0^{x_1} (x_1 + x_2) \, dx_n.$$

26. Demuestre que

$$\int_0^x dx_1 \int_0^{x_1} dx_2 \ldots \int_0^{x_{n-1}} f(x_n) \, dx_n = \frac{1}{(n-1)!} \int_0^x (x - t)^{n-1} f(t) \, dt.$$

2. Integrales múltiples

Las integrales múltiples están estrechamente ligadas a las integrales iteradas de la sección anterior. Supongamos dada una función real definida en un conjunto B en $\mathcal{R}^n$ Nuestro problema es dar una definición de la integral de f en B, análoga a la definición de la integral en una variable, usando sumas en lugar de integrales iteradas.

Primero consideraremos algunos conjuntos simples en $\mathcal{R}^n$. Un **rectángulo coordenado cerrado** es un subconjunto de $\mathcal{R}^n$ formado por todos los puntos $\mathbf{x} = (x_1, \ldots, x_n)$ que satisfacen el conjunto de desigualdades

$$a_i \le x_i \le b_i, \qquad i = 1, \ldots, n. \tag{1}$$

Si en la Ec. (1) algunos de los símbolos «$\le$» son remplazados por «$<$», el conjunto que resulta se llama un **rectángulo coordenado.**

En particular, si todas las desigualdades son de la forma $a_i < x_i < b_i$, el

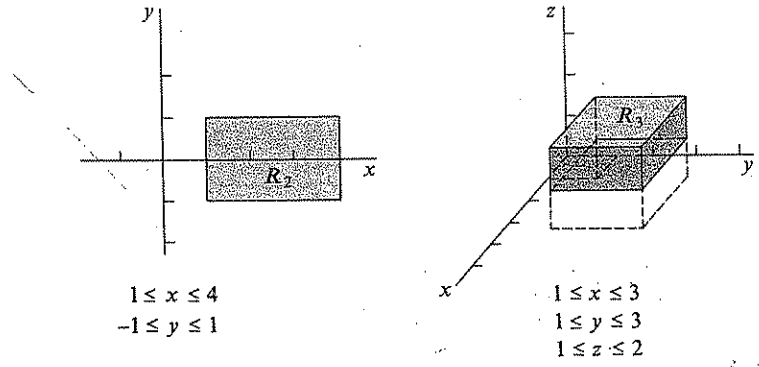


Figura 9

conjunto es abierto y se llama un **rectángulo coordenado abierto.** Un rectángulo coordenado tiene aristas paralelas a los ejes coordenados. En esta sección la palabra «rectángulo» se entenderá como «rectángulo coordenado». La Fig. 9 muestra rectángulos en $\mathcal{R}^2$ y $\mathcal{R}^3$. Un rectángulo en $\mathcal{R}$ es simplemente un intervalo.

Sea R el rectángulo (abierto, cerrado o ninguno de éstos) definido por la Ec. (1) donde se permite remplazar algunos de los símbolos «$\le$» por «$<$». El **volumen** o **contenido** de R, denotado por $V(R)$ es, por definición,

$$V(R) = (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n). \tag{2}$$

En los ejemplos de la Fig. 9, $V(R_2) = (4 - 1)(1 - (-1)) = 6$ y $V(R_3) = (3 - 1)(3 - 1)(2 - 1) = 4$. Si, para algún i en la Ec. (1), $a_i = b_i$, entonces R se llama **degenerado** y $V(R) = 0$. Para rectángulos en $\mathcal{R}^2$, el contenido es lo mismo que el área y a menudo escribimos $A(R)$ en lugar de $V(R)$, para que la notación esté de acuerdo con la noción de área en lugar de la de volumen.

Un subconjunto B de $\mathcal{R}^n$ es **acotado** si existe un número real k tal que $|\mathbf{x}| < k$ para todo $\mathbf{x}$ en B. Un conjunto finito de planos $(n - 1)$-dimensionales en $\mathcal{R}^n$ (rectas en $\mathcal{R}^2$) paralelos a los planos coordenados se llama un **reticulado.** Como lo ilustra la Fig. 10, un reticulado divide a $\mathcal{R}^n$ en un número finito de rectángulos cerrados y acotados $R_1, \ldots, R_r$ y un número finito de regiones no acotadas. Un reticulado **cubre** un subconjunto B de $\mathcal{R}^n$ si B está contenido en la unión de los rectángulos acotados $R_1, \ldots, R_r$.

Es evidente que un conjunto puede ser cubierto por un reticulado si, y solo si, el conjunto es acotado. Como una medida de la finura de un reticulado, tomamos el máximo de los largos de las aristas de los rectángulos $R_1, \ldots, R_r$. Este número se llama la **malla** del reticulado.

Ahora damos una definición de la integral múltiple, llamada también integral de Riemann, en honor de Bernhard Riemann (1826-1866). Consideremos una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ y un conjunto B tal que

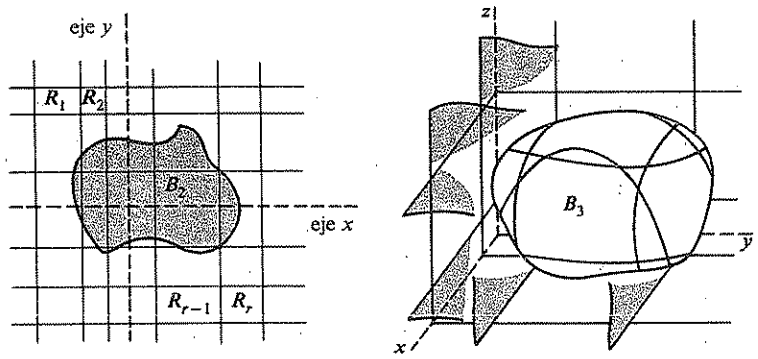


Figura 10

(a) B es un subconjunto acotado del dominio de f.

(b) f es acotada en B.

La afirmación (b) significa que existe un número real K tal que $|f(\mathbf{x})| \leq K$ para todo $\mathbf{x}$ en B. La integral múltiple de f en B será definida en términos de la función f_B que es igual a f siendo cero fuera de B. Es decir,

$$f_B(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{si } \mathbf{x} \text{ está en } B. \\ 0, & \text{si } \mathbf{x} \text{ no está en } B. \end{cases}$$

Sea G un reticulado que cubre a B y con malla igual a $m(G)$. En cada uno de los rectángulos acotados R_i formados por G, $i = 1, \ldots, r$, elegimos un punto arbitrario $\mathbf{x}_i$. La suma

$$\sum_{i=1}^{r} f_B(\mathbf{x}_i) V(R_i)$$

se llama la **suma de Riemann** de f en B. Dados f y B, su valor depende de G y $\mathbf{x}_1, \ldots, \mathbf{x}_r$. Si para todo reticulado G con malla $m(G)$ que tiende a cero, se tiene que

$$\lim_{m(G)\to 0} \sum_{i=1}^{r} f_B(\mathbf{x}_i) V(R_i)$$

existe y es siempre igual al mismo número, entonces este límite es la **integral de f en B** y se denota por $\int_B f\,dV$. Si la integral existe, se dice que f es **integrable en B**.

El límite que define la integral múltiple es algo diferente del límite de funciones vectoriales definido en la Sec. 2 del Cap. 2, a pesar de que la idea es similar. La ecuación

$$\lim_{m(G)\to 0} \sum_{i=1}^{r} f_B(\mathbf{x}_i) V(R_i) = \int_B f\,dV$$

significa que, para todo $\epsilon > 0$, existe $\delta > 0$ tal que si G es un reticulado que cubre a B y que tiene malla menor que δ y S es una suma de Riemann arbitraria para f_B formada con G, entonces

$$\left| S - \int_B f\,dV \right| < \epsilon.$$

Debería recalcarse que la integral no está definida para una función $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ y un conjunto B a no ser que se satisfagan las condiciones de que f y B sean acotadas. Sin estas condiciones, incluso algunas sumas de Riemann pueden no estar definidas.

Si f es una función real de una variable, es decir, si $n = 1$, y B es un intervalo $a \leq x \leq b$, la integral de Riemann de f en B es la integral definida habitualmente.

$$\int_a^b f(x)\,dx.$$

Otras notaciones comunes para la integral de $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ en B son

$$\int_B f\,dA \quad \text{y} \quad \int_B f(x, y)\,dx\,dy, \quad \text{si } n = 2,$$

$$\int_B f(x, y, z)\,dx\,dy\,dz, \quad \text{si } n = 3,$$

$$\int_B f\,dx_1 \ldots dx_n, \quad \text{para } n \text{ arbitrario.}$$

En la mayor parte de las aplicaciones, la integral múltiple puede ser remplazada por la integral iterada, como veremos más tarde. Sin embargo, por razones teóricas es importante tener condiciones bien claras bajo las cuales existe la integral múltiple. El teorema que sigue, cuya demostración es dada en el Apéndice, da tales condiciones.

2.1 Teorema

Sea f definida y acotada en un conjunto acotado B en $\mathcal{R}^n$ tal que la frontera de B está contenida en un número finito de conjuntos lisos. Si f es continua en B excepto quizá en un número finito de conjuntos lisos, entonces f es integrable en B. El valor de la integral $\int_B f\,dV$ no cambia si cambiamos los valores de f en un conjunto liso.

Se entiende por **conjunto liso** en $\mathcal{R}^n$ a la imagen de un conjunto cerrado y acotado por una función diferenciable $\mathcal{R}^m \xrightarrow{\phi} \mathcal{R}^n$, $m < n$. Por tanto, si $n = 2$, y $m = 1$ y podremos tener una curva lisa. Un subconjunto liso de $\mathcal{R}$ es, por definición, un punto. Decir que el valor de $\int_B f\,dV$ es independiente de los valores en un conjunto liso significa que se le pueden asignar valores arbitrarios a f en ese conjunto sin que afecte la existencia o el valor de la integral. Por ejemplo, podemos cambiar el integrando en un conjunto finito de puntos sin

que cambie la integral. Este tipo de modificación es a menudo conveniente para eliminar discontinuidades.

Ejemplo 1. Calculemos la integral múltiple

$$\int_B (2x + y)\, dx\, dy,$$

donde B es el rectángulo $0 \le x < 1$, $0 \le y \le 2$. El Teorema 2.1 garantiza la existencia de esta integral. Por esta razón, cualquier sucesión de sumas de Riemann con mallas que tienden a cero puede ser usada para calcularla. Para

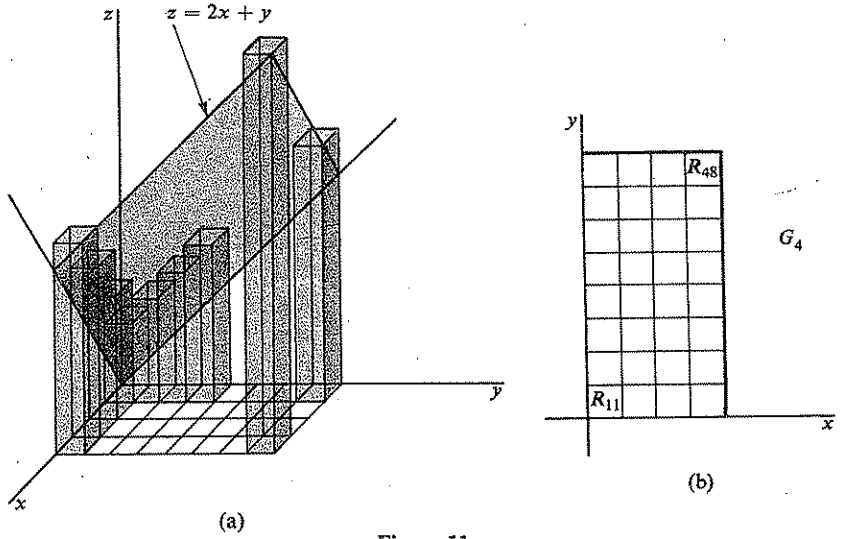


Figura 11

$n = 1, 2, \ldots$, consideremos el reticulado formado por las rectas $x = i/n$, $i = 0, \ldots, n$ e $y = j/n$, $j = 0, \ldots, 2n$. (Véase Fig. 11(b).) La malla de G_n es $1/n$ y el área de cada uno de los rectángulos R_{ij} es $1/n^2$. Usando

$$\mathbf{x}_{ij} = (x_i, y_j) = \left(\frac{i}{n}, \frac{j}{n}\right),$$

formamos la suma de Riemann

$$\sum_{i=1}^{n}\sum_{j=1}^{2n}(2x_i + y_j)A(R_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{2n}\left(\frac{2i}{n} + \frac{j}{n}\right)\frac{1}{n^2}$$

$$= \frac{1}{n^3}\left(4n\sum_{i=1}^{n} i + n\sum_{j=1}^{2n} j\right)$$

$$= \frac{1}{n^2}\left(\frac{4n^2 + n}{2} + \frac{4n^2 + 2n}{2}\right)$$

$$= \frac{4n^2 + 3n}{n^2} = 4 + \frac{3}{n}.$$

mostrada en la Fig. 11(a). Por tanto,

$$\int_B (2x + y)\, dx\, dy = \lim_{n\to\infty}\left(4 + \frac{3}{n}\right) = 4.$$

La evaluación directa de una integral múltiple es, en la mayoría de los casos, muy complicada. Afortunadamente en muchos casos la integral múltiple puede ser calculada aplicando repetidamente la integración 1-dimensional ordinaria en lugar de encontrar los límites de sumas de Riemann. El teorema adecuado, que demostraremos al final de esta sección, es el siguiente:

2.2 Teorema

Sea B un subconjunto de $\mathfrak{R}^n$ tal que la integral iterada

$$\int dx_1 \int dx_2 \ldots \int f\, dx_n$$

existe en B. Si, además, la integral múltiple

$$\int_B f\, dV$$

existe, entonces ambas integrales son iguales.

Como el argumento usado para demostrar el Teorema 2.2 es igualmente válido para cualquier orden de integración, tenemos como corolario inmediato.

2.3 Teorema

Si $\int_B f\, dV$ existe y las integrales iteradas en algunos órdenes existen, entonces todas estas integrales son iguales.

Ejemplo 2. Calculemos la integral $\int_B (2x + y)\, dx\, dy$, donde B es el rectángulo $0 \le x \le 1$, $0 \le y \le 2$. Esta integral es la del Ejemplo 1. Podemos aplicar el Teorema 2.2 y tenemos que

$$\int_B (2x + y)\, dx\, dy = \int_0^1 dx \int_0^2 (2x + y)\, dy$$

$$= \int_0^1 (4x + 2)\, dx$$

$$= [2x^2 + 2x]_0^1 = 4.$$

Ejemplo 3. Sea R el rectángulo tridimensional definido por $-1 \leq x \leq 2$, $0 \leq y \leq 1$, $1 \leq z \leq 2$, mostrado en la Fig. 12. Consideremos $f(x, y, z) = xyz$. Entonces

$$\int_R f\,dV = \int_R xyz\,dx\,dy\,dz$$
$$= \int_{-1}^{2} dx \int_0^1 dy \int_1^2 xyz\,dz = \int_{-1}^{2} x\,dx \int_0^1 y\,dy \int_1^2 z\,dz$$
$$= [(\tfrac{3}{2})(\tfrac{1}{2})(\tfrac{3}{2}) = \tfrac{9}{8}.$$

Ejemplo 4. Sea $f(x, y, z) = xyz$ y sea B el subconjunto de $\mathfrak{R}^3$ definido por $x^2 + y^2 + z^2 \leq 4$, $y \geq 0$, $z \geq 0$. B es el interior y la frontera de un octavo de la bola esférica de radio 2 y centro en el origen. (Véase Fig. 13.) La integral $\int_B f\,dV$ es la integral iterada triple de la función $f(x, y, z) = xyz$ en B. Para x e y fijos, la variable z varía entre 0 y $\sqrt{4 - x^2 - y^2}$ y éstos son los límites de

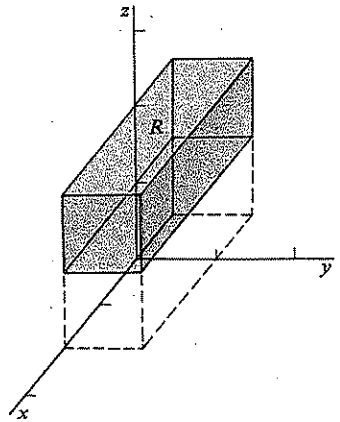


Figura 12

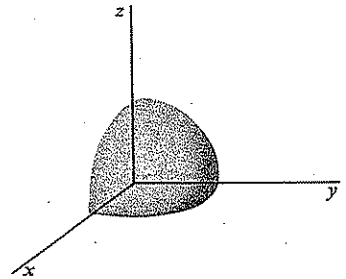


Figura 13

la primera integración con respecto a z. El resultado de esta integración es una función de x e y que debe ser integrada en el subconjunto de dos dimensiones obtenido proyectando B en el plano xy; es decir, en la región $x^2 + y^2 \leq 4$, $y \geq 0$. Para x fijo la variable y varía entre 0 y $\sqrt{4 - x^2}$; por tanto, éstos son los límites de la integración con respecto a y. Por último, x varía entre -2 y 2, de modo que

$$\int_B f\,dV = \int_0^2 dx \int_0^{\sqrt{4-x^2}} dy \int^{\sqrt{4-x^2-y^2}} xyz\,dz.$$

Entonces

$$\int_B f\,dV = \frac{1}{2}\int_0^2 x\,dx \int_0^{\sqrt{4-x^2}} y(4 - x^2 - y^2)\,dy$$

$$= \frac{1}{2}\int_0^2 x\left(2(4 - x^2) - \frac{x^2}{2}(4 - x^2) - \frac{4 - x^2)^2}{4}\right) dx.$$

Esta última integral se reduce a

$$\int_0^2 (2x - x^3 + \tfrac{1}{8}x^5)\,dx = \tfrac{4}{3}.$$

Si la función constante 1 es integrable en un subconjunto B de $\mathfrak{R}^n$, el **contenido** o **volumen** de B es denotado por $V(B)$ y definido por

$$V(B) = \int_B 1\,dV = \int_B dV.$$

Para conjuntos B en $\mathfrak{R}^2$, escribimos $A(B)$, por área, en lugar de $V(B)$. De la última parte del Teorema 2.1 se deduce que el contenido de una curva o superficie S, k-dimensional ($k < n$) y continuamente diferenciable es cero, puesto que

$$V(S) = \int_S dV = 0.$$

Para algunos conjuntos B, la integral $\int_B dV$ no existe. En este caso, el contenido de B no está definido (véase Ejercicio 21). Nótese que si R es un rectángulo, $V(R)$ ha sido definido dos veces: primero como el producto de los largos de aristas enteramente perpendiculares y segundo como una integral. De los Teoremas 2.1 y 2.2 se deduce inmediatamente que ambas definiciones coinciden.

Ejemplo 5. Sea B la región en $\mathfrak{R}^2$ bajo la curva $y = f(x)$ desde $x = a$ hasta $x = b$, donde f es una función no negativa. Si suponemos que existen las integrales consideradas, usando el teorema de la integral iterada (Teorema 2.2), tenemos que

$$A(B) = \int_B dA = \int_a^b dx \int_0^{f(x)} dy = \int_a^b f(x)\,dx.$$

Por tanto, la definición de contenido que hemos dado anteriormente está de acuerdo con la definición habitual de área bajo el gráfico de una función no negativa e integrable de una variable. Si f es integrable y no negativa en B, podríamos demostrar en forma análoga que el volumen bajo el gráfico de f y sobre B es la integral doble $\int_B f\,dA$.

Ejemplo 6. El volumen sobre el disco D definido por $x^2 + y^2 \leq 1$ y bajo el gráfico de $f(x, y) = x^2 + y^2$ (véase Fig. 14) es igual a

$$\int_D (x^2 + y^2)\,dx\,dy = \int_{-1}^{1} dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2 + y^2)\,dy$$
$$= 2\int_{-1}^{1} (x^2\sqrt{1 - x^2} + \tfrac{1}{3}(1 - x^2)\sqrt{1 - x^2})\,dx$$

$$= \frac{4}{3}\int_0^1 (\sqrt{1-x^2} + 2x^2\sqrt{1-x^2})\,dx$$

$$= \frac{4}{3}\left(\frac{\pi}{4} + \frac{\pi}{8}\right) = \frac{\pi}{2}.$$

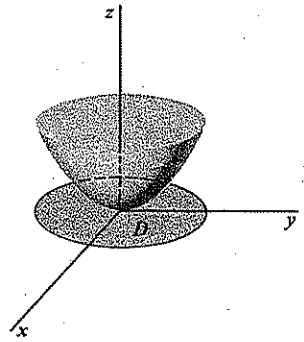


Figura 14

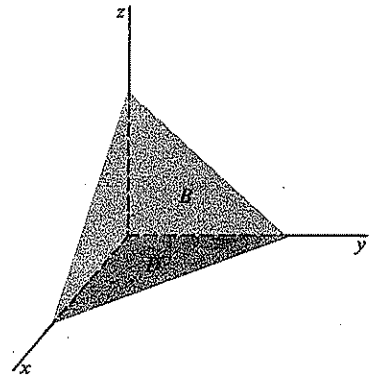


Figura 15

Ejemplo 7. Encontraremos el volumen de la región B en $\mathfrak{R}^3$ acotada por los cuatro planos $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$, mostrado en la Fig. 15.

$$V(B) = \int_B dV = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz = \tfrac{1}{6}.$$

El volumen de la región B puede ser calculado directamente como una integral doble. La proyección de B en el plano xy es el triángulo D acotado por las rectas $x = 0$, $y = 0$, $x + y = 1$. El conjunto B puede ser descrito como la región bajo el gráfico de la función $f(x, y) = 1 - x - y$ que está sobre D. Por tanto, de acuerdo con la afirmación hecha al final del Ejemplo 5,

$$V(B) = \int_D f\,dA = \int_0^1 dx \int_0^{1-x} (1 - x - y)\,dy = \tfrac{1}{6}.$$

Nótese que lo que hemos definido como contenido de un subconjunto de $\mathfrak{R}^n$ debería ser llamado **contenido n-dimensional**. Por ejemplo, el cuadrado definido por las desigualdades $0 \le x \le 2$, $0 \le y \le 2$ en $\mathfrak{R}^2$ tiene contenido 2-dimensional 4, mientras que el cuadrado en $\mathfrak{R}^3$ definido por $0 \le x \le 2$, $0 \le y \le 2$, $z = 0$ y que tiene la misma forma, tiene contenido 3-dimensional 0. En otras palabras, el contenido de un conjunto depende de la dimensión del espacio euclidiano que lo contiene y con respecto del cual se lo está midiendo. Habiendo indicado que el contenido 2-dimensional se llama área, anotamos que el contenido 1-dimensional es el largo.

En los cuatro teoremas que siguen resumimos algunas propiedades características de la integral de Riemann.

2.4 Teorema. Linealidad

Si f y g son integrables en B y a y b son dos números reales, entonces $af + bg$ es integrable en B y

$$\int_B (af + bg)\,dV = a\int_B f\,dV + b\int_B g\,dV.$$

2.5 Teorema. Positividad

Si f es no negativa e integrable en B, entonces

$$\int_B f\,dV \ge 0.$$

2.6 Teorema

Si R es un rectángulo, entonces $\int_R dV = V(R)$ (donde el contenido $V(R)$ está definido por la Ec. (2)).

2.7 Teorema

Si B es un subconjunto de un conjunto acotado C, entonces $\int_B f\,dV$ existe si, y solo si, $\int_C f_B\,dV$ existe. Cuando las dos integrales existen, ellas son iguales.

Demostración de 2.4. Sea $\epsilon > 0$ dado y elijamos $\delta > 0$ tal que si S_1 y S_2 son dos sumas de Riemann de f_B y g_B, respectivamente, cuyos reticulados tienen malla menor que δ, entonces

$$|a|\left|S_1 - \int_B f\,dV\right| < \frac{\epsilon}{2} \quad \text{y} \quad |b|\left|S_2 - \int_B g\,dV\right| < \frac{\epsilon}{2}.$$

Sea S cualquier suma de Riemann de $(af + bg)_B$ cuyo reticulado tiene malla menor que δ. Entonces

$$S = \sum_i (af + bg)_B(\mathbf{x}_i)V(R_i)$$

$$= a\sum_i f_B(\mathbf{x}_i)V(R_i) + b\sum_i g_B(\mathbf{x}_i)V(R_i)$$

$$= aS_1 + bS_2.$$

Por tanto,

$$\left|S - a\int_B f\,dV - b\int_B g\,dV\right|$$

$$= \left| aS_1 - a\int_B f\,dV + bS_2 - b\int_B g\,dV \right|$$

$$\leq |a| \left| S_1 - \int_B f\,dV \right| + |b| \left| S_2 - \int_B g\,dV \right|$$

$$< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

luego

$$\lim_{m(G)\to 0} \sum_i (af + bg)_B(\mathbf{x}_i)V(R_i) = a\int_B f\,dV + b\int_B g\,dV,$$

y la demostración queda terminada.

Demostración de 2.5. Como todas las sumas de Riemann son no negativas, su límite también debe ser no negativo.

Demostración de 2.6. Inmediata, usando los Teoremas 2.1 y 2.2.

Demostración de 2.7. La existencia y el valor de la integral $\int_B f\,dV$ dependen sólo de la función f_B. Análogamente, $\int_C f_B\,dV$ se define usando $(f_B)_C$ que es igual a f_B.

Muchas propiedades importantes de la integral se deducen directamente de los cuatro teoremas anteriores sin recurrir a la definición original. Como ejemplo, tenemos los dos teoremas siguientes.

2.8 Teorema

Si f y g son integrables en B y $f \leq g$, entonces

$$\int_B f\,dV \leq \int_B g\,dV.$$

Demostración. La función $g - f$ es no negativa y, por el Teorema 2.4, es integrable en B. Por tanto, por los Teoremas 2.4 y 2.5,

$$0 \leq \int_B (g - f)\,dV = \int_B g\,dV - \int_B f\,dV,$$

de donde se tiene el resultado buscado.

El segundo teorema establece una propiedad análoga a la ecuación

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

válida para funciones de una variable.

2.9 Teorema

Si f es integrable en cada uno de los conjuntos disjuntos B_1 y B_2, entonces f es integrable en su unión y

$$\int_{B_1\cup B} f\,dV = \int_{B_1} f\,dV + \int_{B_2} f\,dV.$$

Demostración. Por el Teorema 2.7,

$$\int_{B_1} f\,dV + \int_{B_2} f\,dV = \int_{B_1\cup B_2} f_{B_1}\,dV + \int_{B_1\cup B_2} f_{B_2}\,dV.$$

Como B_1 y B_2 son disjuntos, $f_{B_1\cup B_2} = f_{B_1} + f_{B_2}$. Luego, por el Teorema 2.4, $f_{B_1\cup B_2}$ es integrable en $B_1 \cup B_2$ y

$$\int_{B_1\cup B_2} f_{B_1}\,dV + \int_{B_1\cup B_2} f_{B_2}\,dV = \int_{B_1\cup B_2} f_{B_1\cup B_2}\,dV.$$

Por último, usando nuevamente el Teorema 2.7, f es integrable en $B_1 \cup B_2$ y

$$\int_{B_1\cup B_2} f_{B_1\cup B_2}\,dV = \int_{B_1\cup B_2} f\,dV.$$

Esto completa la demostración.

El próximo teorema mostrará que, para funciones f y regiones B para las cuales $\int_B f\,dV$ existe, el valor de la integral está completamente determinado por las propiedades dadas en los Teoremas 2.4-2.7. El teorema es importante, ya que nos permite identificar las integrales múltiples con otras integrales I, por ejemplo, con las integrales iteradas que pueden ser calculadas usando cualquier orden. Usaremos el símbolo I para designar una integral de este tipo.

2.10 Teorema

Sea I una función que asigna un número real $I_B f$ a ciertas funciones acotadas $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}$ y conjuntos acotados B. Supongamos que I satisface las siguientes propiedades:

(a) Si $I_B f$ e $I_B g$ están definidas y a y b son números reales, entonces $I_B(af + bg)$ está definida y

$$I_B(af + bg) = aI_B f + bI_B g.$$

(b) Si f es no negativa e $I_B f$ está definida, entonces $I_B f \geq 0$.

(c) Si R es un rectángulo, entonces $I_R 1 = V(R)$ (como ha sido definida por la Ec. (2)).

(d) Si B está contenido en un conjunto acotado C, entonces $I_B f$ está definido si, y solo si, $I_C f_B$ está definido. Cuando ambos existen, ellos son iguales.

Concluimos que si $I_B f$ y $\int_B f\,dV$ existen, ellos son iguales.

Demostración. Supongamos que $\int_B f\,dV < I_B f$. Hagamos

$$\epsilon = I_B f - \int_B f\,dV,$$

y elijamos $\delta > 0$ tal que si S es una suma de Riemann de f_B cuyo reticulado tiene malla menor que δ, entonces

$$\left| \int_B f\,dV - S \right| < \frac{\epsilon}{2}.$$

Sea G un reticulado arbitrario que cubre a B y cuya malla es menor que δ y denotemos por $R_1, \ldots, R_r$ los rectángulos formados por G. Sea

$$C = R_1 \cup \ldots \cup R_r,$$

$$f_i = \text{ínfimo de } f_B \text{ en } R_i.$$

Consideremos la función g definida por

$$g = \sum_{i=1}^{r} f_i \chi_{R_i}.$$

La función χ_{R_i} es la función característica de R_i. Está definida por

$$\chi_{R_i}(\mathbf{x}) = \begin{cases} 1, & \text{si x está en } R_i, \\ 0, & \text{en caso contrario.} \end{cases}$$

De (c), (d) y (a) se deduce inmediatamente que $I_C g$ está definida y que

$$I_C g = \sum_{i=1}^{r} f_i V(R_i).$$

La definición del ínfimo implica que existe una suma de Riemann para f_B en el reticulado G que está arbitrariamente cerca de $I_C g$. Por tanto,

$$\left| \int_B f\,dV - I_C g \right| \leq \frac{\epsilon}{2},$$

y, por definición de ϵ,

$$I_C g < I_B f. \tag{3}$$

Por (d), $I_B f = I_C f_B$. Además, la función g ha sido construida de modo que $f_B \leq g$. Por tanto, de (b) (extendido como en el Teorema 2.8) se tiene que

$$I_B f = I_C f_B \leq I_B g. \tag{4}$$

Las desigualdades (3) y (4) son contradictorias, de modo que concluimos que

$$\int_B f\,dV \geq I_B f.$$

Mediante un argumento totalmente análogo, usando la noción de supremo en lugar de ínfimo, obtenemos que

$$\int_B f\,dV \leq I_B f,$$

y esto termina la demostración.

Como aplicación del Teorema 2.10, tomemos las funciones $\mathfrak{R}^2 \xrightarrow{f} \mathfrak{R}$ y

los conjuntos B para los cuales la integral iterada $\int dx \int f\,dy$ en B está definida. Sea

$$I_B f = \int dx \int_{\text{(sobre } B\text{)}} f\,dy.$$

La verificación de las condiciones del Teorema 2.10 no tiene dificultades y se reduce a las propiedades correspondientes para la integral definida de funciones de una variable. Por ejemplo, para la integral en intervalos,

$$\int_\alpha^\beta (af + bg)\,dx = a\int_\alpha^\beta f\,dx + b\int_\alpha^\beta g\,dx.$$

$$\int_\alpha^\beta f\,dx \geq 0, \qquad \text{si } f \geq 0.$$

$$\int_\alpha^\beta dx = \beta - \alpha.$$

$$\int_\gamma^\delta f\,dx = \int_\alpha^\beta f_{(\gamma,\delta)}\,dx,$$

si $\alpha \leq \gamma \leq \delta \leq \beta$ y $[\gamma, \delta]$ es el intervalo $\gamma \leq x \leq \delta$.

Del Teorema 2.10 se deduce de inmediato que si la integral iterada y la integral doble de f existen en B, ellas son iguales. Esto demuestra el Teorema 2.2 para dos variables. El caso general se hace por inducción.

La posibilidad de intercambiar el orden de integración tiene otras consecuencias además de la conveniencia para la computación de integrales múltiples. Una de éstas es el teorema sobre el cambio de orden en la diferenciación parcial demostrada por otros métodos en la Sec. 5 del Cap. 3 y dado en una forma ligeramente más fuerte en el Ejercicio 8. Otra consecuencia es la fórmula de Leibnitz para intercambiar el orden de diferenciación e integración. Indicamos a continuación el teorema y su demostración es indicada en el Ejercicio 7.

2.11 Regla de Leibnitz

Si $(\partial g/\partial y)(x, y)$ es continua en $a \leq x \leq b$ y $c \leq y \leq d$, entonces

$$\frac{d}{dy}\int_a^b g(x, y)\,dx = \int_a^b \frac{\partial g}{\partial y}(x, y)\,dx.$$

EJERCICIOS

1. Haga un dibujo del conjunto B y calcule $\int_B f\,dA$, donde

   (a) $f(x, y) = x^2 + 3y^2$ y B es el disco $x^2 + y^2 \leq 1$. [*Resp.* π.]

   (b) $f(x, y) = 1/(x + y)$ y B es la región acotada por las rectas $y = x$, $x = 1$, $x = 2$, $y = 0$. [*Resp.* $\log 2$.]

(c) $f(x, y) = x \operatorname{sen} xy$ y B es el rectángulo $0 \le x \le \pi$, $0 \le y \le 1$. [*Resp.* π.]

(d) $f(x, y) = x^2 - y^2$ y B está formado por todos los (x, y) tales que $0 \le x \le 1$ y $x^2 - y^2 \ge 0$. [*Resp.* $\frac{1}{3}$.]

2. Usando la definición de la integral doble como un límite de sumas de Riemann, calcule $\int_B f(x, y)\, dx\, dy$, donde

(a) $f(x, y) = x + 4y$ y B es el rectángulo $0 \le x \le 2$, $0 \le y \le 1$. [*Resp.* 6.]

(b) $f(x, y) = 3x^2 + 2y$ y B es el rectángulo $0 \le x \le 2$, $0 \le y \le 1$. [*Resp.* 10.]

Las siguientes fórmulas serán útiles para resolver este problema:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2},$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6},$$

$$\sum_{i=1}^{n} i^3 = \left(\sum_{i=1}^{n} i\right)^2.$$

3. Encuentre el volumen bajo el gráfico de f y sobre el conjunto B, donde

(a) $f(x, y) = x + y^2$ y B es el rectángulo con vértices (1, 1), (1, 3), (2, 3) y (2, 1). [*Resp.* $\frac{35}{3}$.]

(b) $f(x, y) = x + y + 2$ y B es la región acotada por las curvas $y^2 = x$ y $x = 2$. [*Resp.* $\frac{128}{15}\sqrt{2}$.]

(c) $f(x, y) = |x + y|$ y B es el disco $x^2 + y^2 \le 1$. [*Resp.* $4\sqrt{2}/3$.]

4. Encuentre por integración el área del subconjunto de $\mathcal{R}^2$ acotado por la curva $x^2 - 2x + 4y^2 - 8y + 1 = 0$. [*Resp.* 2π.]

5. Encuentre un valor aproximado para cada una de las integrales del Ejercicio 1, calculando una suma de Riemann con reticulado apropiadamente fino.

6. Considere los rectángulos

B_1 definido por $0 \le x \le 1$, $0 \le y \le 1$
B_2 definido por $1 \le x \le 2$, $-1 \le y \le 1$

y la función

$$f(x, y) = \begin{cases} 2x - y, & \text{si } x < 1, \\ x^2 + y, & \text{si } x \ge 1. \end{cases}$$

Calcule $\int_{B_1 \cup B_2} f(x, y)\, dx\, dy$. [*Resp.* $\frac{31}{6}$.]

7. (a) Demuestre la **regla de Leibnitz** para diferenciar una integral con respecto a un parámetro. Si $g_y(x, y)$ es continua en un rectángulo $a \le x \le b$, $c \le y \le d$, entonces

$$\frac{d}{dy}\int_a^b g(t, y)\, dt = \int_a^b g_y(t, y)\, dt.$$

[*Indicación*: Cambie el orden de integración en $\int_c^y dy \int_a^b g_y(t, y)\, dt$, y después derive ambos términos con respecto a y.]

(b) Use la parte (a) y la regla de la cadena para demostrar que si g_y es continua y h_1 y h_2 son diferenciables, entonces

$$\frac{d}{dy}\int_{h_1(y)}^{h_2(y)} g(t, y)\, dt = \int_{h_1(y)}^{h_2(y)} g_y(t, y)\, dt + h_2'(y)g(h_2(y), y) - h_1'(y)g(h_1(y), y).$$

8. Demuestre (compare con 5.2 del Cap. 3): Si f_x, f_y y f_{xy} son continuas en un conjunto abierto, entonces $f_{xy} = f_{yx}$. [*Indicación*: Aplique la regla de Leibnitz del Ejercicio 7(a) a la ecuación

$$f(x, y) - f(a, y) = \int_a^x f_x(t, y)\, dt,$$

y luego derive ambos lados con respecto a x.]

9. Si $f(x, y, z) = xyz$ y

$$\int_B f(x, y, z)\, dx\, dy\, dz = \int_0^2 dx \int_0^x dy \int_0^{x+y} xyz\, dz,$$

dibuje la región B y calcule la integral. [*Resp*: $\frac{68}{9}$.]

10. Calcule la integral múltiple de $f(x, y, z, w) = xyzw$ en el rectángulo 4-dimensional

$$0 \le x \le 1, \quad -1 \le y \le 2, \quad 1 \le z \le 2, \quad 2 \le w \le 3. \quad [\textit{Resp.}\ \tfrac{45}{16}.]$$

11. Dibuje la región B en $\mathcal{R}^3$ acotada por la superficie $z = 4 - 4x^2 - y^2$ y el plano xy. Exprese el volumen de B como una integral triple y como una integral doble. Calcule el volumen. [*Resp.* 4π.]

12. Escriba una expresión para el volumen de la bola $x^2 + y^2 + z^2 \le a^2$,

(a) como una integral triple,
(b) como una integral doble.

13. Dibuje en $\mathcal{R}^3$ los dos cilindros sólidos definidos por $x^2 + z^2 \le 1$ e $y^2 + z^2 \le 1$. Encuentre el volumen de la intersección. [*Resp.* $\frac{16}{3}$.]

14. La bola 4-dimensional B de radio 1 y centro en el origen es el subconjunto de $\mathcal{R}^4$ definido por $x_1^2 + x_2^2 + x_3^2 + x_4^2 \le 1$. Encuentre una expresión para el volumen $V(B)$ como una integral cuádruple.

15. Use el Teorema 2.8 para demostrar que si f y $|f|$ son integrables en B, entonces $\left|\int_B f\, dV\right| \ge \int_B |f|\, dV$.

16. Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ definida en un conjunto B de $\mathcal{R}^n$. Definimos

$$\int_B f\, dV = \left(\int_B f_1\, dV, \ldots, \int_B f_m\, dV\right),$$

siempre que existan las integrales de las funciones coordenadas $f_1, \ldots, f_m$ de f.

(a) Demuestre que si $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ y $\mathfrak{R}^n \xrightarrow{g} \mathfrak{R}^m$ son integrables en B, entonces

$$\int_B (af + bg)\,dV = a\int_B f\,dV + b\int_B g\,dV,$$

donde a y b son constantes.

(b) Si $\mathbf{k}$ es un vector constante en $\mathfrak{R}^m$ y $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ es integrable en B, demuestre que

$$\int_B \mathbf{k}\cdot f\,dV = \mathbf{k}\cdot\int_B f\,dV.$$

(c) Demuestre que si $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ y $\mathfrak{R}^n \xrightarrow{|f|} \mathfrak{R}$ son integrables en B, entonces $|\int_B f\,dV| \le \int_B |f|\,dV$. [*Indicación*: Por la desigualdad de Cauchy-Schwarz $f(\mathbf{x})\cdot\int_B f\,dV \le |f(\mathbf{x})||\int_B f\,dV|$ para todo $\mathbf{x}$ en B. Integre con respecto a $\mathbf{x}$ y aplique el resultado de la parte (b).]

17. Use el resultado del Ejercicio 16(c) para demostrar que si $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}^m$ es continua en un conjunto B y $\mathbf{x}_0$ es un punto interior de B, entonces

$$\lim_{r\to 0}\frac{1}{V(B_r)}\int_{B_r} f\,dV = f(\mathbf{x}_0),$$

donde B_r es la bola de radio r y centro $\mathbf{x}_0$.

18. Sea R una región en $\mathfrak{R}^n$ con volumen finito. El vector

$$\mathbf{z}_0 = \frac{1}{V(R)}\int_R \mathbf{x}\,dV$$

se llama el **centroide** de R y el número real

$$I(\mathbf{z}) = \int_R |\mathbf{x}-\mathbf{z}|^2\,dV_{\mathbf{x}}$$

el **momento de inercia** de R con respecto a $\mathbf{z}$.

(a) Demuestre que el centroide de una bola es su centro.
(b) Demuestre que $I(\mathbf{z})$ tiene su mínimo tomando como $\mathbf{z}$ al centroide de R. [*Indicación*: Demuestre que

$$I(\mathbf{z}) = I(0) - V(R)\,|\mathbf{z}_0|^2 + V(R)\,|\mathbf{z}-\mathbf{z}_0|^2.]$$

19. Demuestre la siguiente propiedad de las integrales múltiples, análoga al Teorema 8.4 del Cap. 5: Si R tiene volumen finito y la serie $\sum_{k=1}^{\infty} f_k$ de funciones continuas converge uniformemente a f en R, entonces

$$\sum_{k=1}^{\infty}\int_R f_k\,dV = \int_R\left[\sum_{k=1}^{\infty} f_k\right]dV.$$

20. Una función $\mathfrak{R}^n \xrightarrow{p} \mathfrak{R}$ integrable en una región R en $\mathfrak{R}^n$ se llama una **densidad de probabilidad** en R si

$$p(\mathbf{x}) \ge 0 \text{ para todo } \mathbf{x} \text{ en } R. \tag{1}$$

$$\int_R p\,dV = 1. \tag{2}$$

Si E es un experimento con resultados posibles en $\mathfrak{R}^n$, distribuidos de acuerdo con la densidad p, entonces la **probabilidad** que el resultado esté en un conjunto B en $\mathfrak{R}^n$ está definida por

$$Pr[E \text{ en } B] = \int_B p\,dV.$$

(a) ¿Para qué constante k, la función $\mathfrak{R}^2 \xrightarrow{p} \mathfrak{R}$,

$$p(x,y) = \begin{cases} k(1-x^2-y^2), & x^2+y^2 \le 1 \\ 0, & x^2+y^2 > 1, \end{cases}$$

es una densidad de probabilidad? [*Resp.* $k = 2/\pi$.]

(b) Si los resultados de E están distribuidos de acuerdo con la densidad de la parte (a), encuentre la probabilidad de que E tenga una coordenada x mayor que $\frac{1}{2}$.

21. Sea B el subconjunto de $\mathfrak{R}^2$ formado por todos los puntos (x, y) tales que $0 \le y \le 1$ y x es racional, $0 \le x \le 1$. ¿Cuál es el área de B?

22. Sea $f(x, y) = 1$ si x es racional, $f(x, y) = 2y$ si x es irracional, para todo (x, y) en el rectángulo $0 \le x \le 1$, $0 \le y \le 1$. Demuestre que

$$\int_0^1 dx\int_0^1 f(x,y)\,dy = 1,$$

y que f no es integrable, en el sentido de Riemann, en el rectángulo.

23. Demuestre que

$$\int_0^1 dy\int_1^\infty (e^{-xy} - 2e^{-2xy})\,dx \ne \int_1^\infty dx\int_0^1 (e^{-xy} - 2e^{-2xy})\,dy.$$

3. Cambio de variable

La fórmula para el cambio de variables de integrales 1-dimensionales es

$$\int_{\phi(a)}^{\phi(b)} f(x)\,dx = \int_a^b f(\phi(u))\phi'(u)\,du. \tag{1}$$

Por ejemplo, tomando $\phi(u) = \operatorname{sen} u$, tenemos

$$\int_0^1 \sqrt{1-x^2}\,dx = \int_0^{\pi/2} \cos^2 u\,du = \frac{\pi}{4}.$$

En esta sección extenderemos la Ec. (1) a dimensiones mayores que 1. En el espacio n-dimensional, se efectúa un cambio de variables mediante una función $\mathfrak{R}^n \xrightarrow{T} \mathfrak{R}^n$. En lo que sigue, generalmente será más conveniente considerar que el espacio dominio y el espacio imagen de T son distintos. Por tanto, consideramos T como una transformación de una copia de $\mathfrak{R}^n$, que designamos por $\mathfrak{U}^n$, en otra copia, que seguimos designando por $\mathfrak{R}^n$, y escribiremos $T(\mathbf{u}) = \mathbf{x}$, donde $\mathbf{u}$ está en $\mathfrak{U}^n$ y $\mathbf{x}$ en $\mathfrak{R}^n$. El enunciado del teorema del cambio de variables en n dimensiones es el siguiente:

3.1 Teorema

$\mathfrak{U}^n \xrightarrow{T} \mathfrak{R}^n$ una transformación continuamente diferenciable. Sea R un conjunto en $\mathfrak{U}^n$ cuya frontera está formada por un número finito de conjuntos lisos. Supongamos que R y su frontera están contenidos en el interior del dominio de T y que

(a) T es inyectiva en R.
(b) det T', el determinante jacobiano de T, es distinto de cero en R.

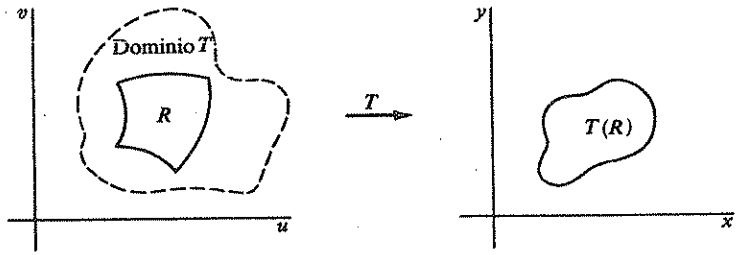


Figura 16

Entonces, si la función f es acotada y continua en $T(R)$ (la imagen de R por T), entonces

$$\int_{T(R)} f\,dV = \int_R (f \circ T)\,|\det T'|\,dV.$$

Es permisible que (a) o (b) sean falsas en un conjunto de contenido cero.

La demostración se dará en el Apéndice. Antes de mostrar cómo funciona esta fórmula, daremos algunos ejemplos de su aplicación. Nótese que el factor ϕ' que aparece en la Ec. (1) ha sido remplazado en dimensiones mayores, por el valor absoluto del determinante jacobiano de T. Aparte del cálculo de det T', la aplicación de la fórmula de transformación consiste en encontrar la relación geométrica entre el subconjunto R y su imagen $T(R)$, para distintas transformaciones T.

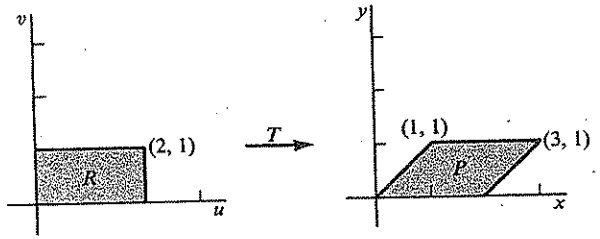


Figura 17

Ejemplo 1. La integral $\int_P (x + y)\,dx\,dy$, en la cual P es el paralelogramo mostrado en la Fig. 17 puede ser transformado en una integral sobre un rectángulo. Esto se hace mediante la transformación

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ v \end{pmatrix}$$

El determinante jacobiano de T es

$$\det T' = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1.$$

Por el teorema del cambio de variables,

$$\int_P (x + y)\,dx\,dy = \int_R [(u + v) + v]1\,du\,dv$$
$$= \int_0^2 du \int_0^1 (u + 2v)\,dv = 4.$$

Es claro que la transformación T es inyectiva, puesto que es una transformación lineal con determinante distinto de cero. Nótese que la región de integración en la integral dada es la imagen de la transformación y no su dominio.

Ejemplo 2. La transformación

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \end{pmatrix}$$

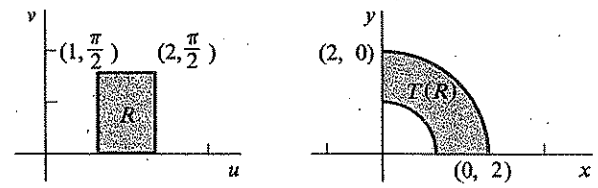


Figura 18

hace corresponder las regiones de la Fig. 18. El jacobiano es

$$\det T' = \begin{vmatrix} \cos v & -u \operatorname{sen} v \\ \operatorname{sen} v & u \cos v \end{vmatrix} = u.$$

La transformación es biunívoca entre R y $T(R)$. Esto se puede ver geométricamente por la interpretación de u y v como el ángulo y el radio respectivamente o directamente de las relaciones

$$u = \sqrt{x^2 + y^2}, \qquad \cos v = \frac{x}{\sqrt{x^2 + y^2}},$$

y del hecho que $\cos v$ es inyectiva para $0 \leq v \leq \pi/2$. La integral de $x^2 + y^2$ en $T(R)$ puede ser transformada como sigue:

$$\int_{T(R)} (x^2 + y^2)\, dA = \int_R u^2 u\, dA = \int_1^2 u^3\, du \int_0^{\pi/2} dv = \frac{15\pi}{8}.$$

Ejemplo 3. Sea B el subconjunto del espacio tridimensional $\mathcal{R}^3$ definido por las desigualdades

$$x^2 + y^2 + z^2 \leq 1, \qquad \begin{cases} x \geq 0, \\ y \geq 0, \\ z \geq 0. \end{cases}$$

Para transformar la integral $\int_B (x^2 + y^2)\, dx\, dy\, dz$, podemos definir T por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = T\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u \operatorname{sen} v \cos w \\ u \operatorname{sen} v \operatorname{sen} w \\ u \cos v \end{pmatrix}$$

Restringiendo (u, v, w) al rectángulo definido por

$$\begin{cases} 0 \leq u \leq 1, \\ 0 \leq v \leq \pi/2, \\ 0 \leq w \leq \pi/2, \end{cases}$$

tenemos que $T(R) = B$. Las regiones correspondientes están representadas en la Fig. 19. Como

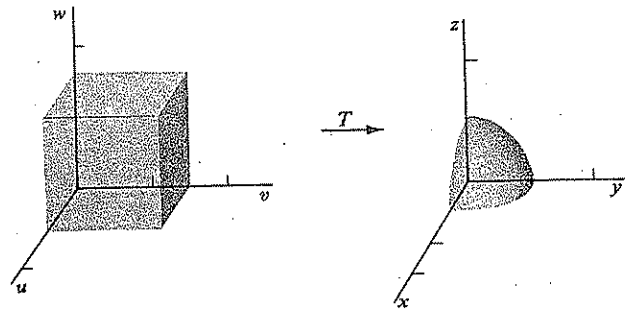


Figura 19

$$u = \sqrt{x^2 + y^2 + z^2}$$

$$\cos v = \frac{z}{\sqrt{x^2 + y^2 + z^2}},$$

$$\cos w = \frac{x}{\sqrt{x^2 + y^2}},$$

deducimos que la transformación T es inyectiva de R en B excepto en los planos $u = 0$ y $v = 0$ en la frontera. El determinante jacobiano es

$$\det T' = \begin{vmatrix} \operatorname{sen} v \cos w & u \cos v \cos w & -u \operatorname{sen} v \operatorname{sen} w \\ \operatorname{sen} v \operatorname{sen} w & u \cos v \operatorname{sen} w & u \operatorname{sen} v \cos w \\ \cos v & -u \operatorname{sen} v & 0 \end{vmatrix}$$

$$= u^2 \operatorname{sen} v.$$

La integral transformada es

$$\int_B (x^2 + y^2)\, dx\, dy\, dz = \int_R (u^2 \operatorname{sen}^2 v \cos^2 w + u^2 \operatorname{sen}^2 v \operatorname{sen}^2 w) u^2 \operatorname{sen} v\, du\, dv\, dw$$

$$= \int_0^1 u^4\, du \int_0^{\pi/2} \operatorname{sen}^3 v\, dv \int_0^{\pi/2} dw$$

$$= \frac{1}{5} \cdot \frac{2}{3} \cdot \frac{\pi}{2} = \frac{\pi}{15}.$$

Nótese que en la integral transformada aparecen valores de u y v para los cuales el jacobiano $u^2 \operatorname{sen} v$ es cero; es decir, $u = 0$ y $v = 0$. Hemos hecho notar que la transformación no es inyectiva en los puntos que satisfacen estas condiciones. Sin embargo, el conjunto de los puntos en que esto sucede, tiene contenido cero, de modo que se puede aplicar el teorema del cambio de variables. Por supuesto que no se afecta el valor de ninguna de las dos integrales si se incluyen o no estos puntos.

Ejemplo 4. Sea la función $\mathcal{U}^2 \xrightarrow{T} \mathcal{R}^2$ definida por

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 - v \\ u + v^2 \end{pmatrix}$$

El cuadrado unitario R_{uv} definido por las desigualdades $0 \leq u \leq 1$, $0 \leq v \leq 1$ es llevado por T sobre el subconjunto R_{xy} mostrado en la Fig. 20. En esta figura se indican las porciones correspondientes de las fronteras. La imagen de cada uno de los cuatro segmentos que forman la frontera de R_{uv} son los siguientes:

(a) Si $u = 0$ y $0 \leq v \leq 1$, entonces $x = -v$ e $y = v^2$. Es decir, $y = x^2$ y $-1 \leq x \leq 0$.

(b) Si $v = 0$ y $0 \leq u \leq 1$, entonces $x = u^2$ e $y = u$. Es decir, $x = y^2$ y $0 \leq y \leq 1$.

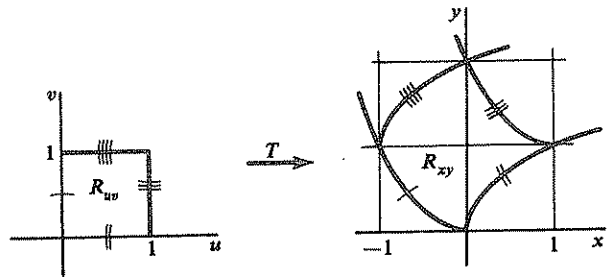


Figura 20

(c) Si $u = 1$ y $0 \le v \le 1$, entonces $x = 1 - v$ e $y = 1 + v^2$. Es decir, $y - 1 = (x - 1)^2$ y $0 \le x \le 1$.

(d) Si $v = 1$ y $0 \le u \le 1$, entonces $x = u^2 - 1$ e $y = u + 1$. Es decir, $(y - 1)^2 = x + 1$ y $1 \le y \le 2$.

No es difícil verificar que T es inyectiva en R_{uv}. En efecto, si suponemos que

$$T\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = T\begin{pmatrix} u_2 \\ v_2 \end{pmatrix},$$

entonces

$$u_1^2 - v_1 = u_2^2 - v_2,$$
$$u_1 + v_1^2 = u_2 + v_2^2.$$

Es claro que si $u_1 = u_2$, entonces $v_1 = v_2$. Supongamos $u_1 < u_2$; entonces

$$0 < u_2^2 - u_1^2 = v_2 - v_1,$$
$$0 < u_2 - u_1 = v_1^2 - v_2^2.$$

Por tanto, $v_1 < v_2$ mientras que $v_2^2 < v_1^2$. Esto es imposible si v_1 y v_2 son no negativos, de modo que la inyectividad de T en R_{uv} está demostrada. El determinante jacobiano de T es

$$\det T' = \begin{vmatrix} 2u & -1 \\ 1 & 2v \end{vmatrix} = 4uv + 1.$$

y, como aplicación del teorema del cambio de variable, tenemos

$$\int_{Rxy} x\,dx\,dy = \int_{Ruv} (u^2 - v)(4uv + 1)\,du\,dv$$
$$= \int_0^1 dv \int_0^1 (4u^3v - 4uv + u^2 - v)\,du$$
$$= \int_0^1 (-2v^2 + \tfrac{1}{3})\,dv = -\tfrac{1}{3}.$$

Para entender por qué es válida la fórmula del cambio de variables para una transformación vectorial diferenciable, necesitamos conocer el efecto de T

en los volúmenes. Usamos la aproximación afín que remplaza $T(\mathbf{u})$ cerca de $\mathbf{u}_0$ por $T(\mathbf{u}_0) + T'(\mathbf{u}_0)(\mathbf{u} - \mathbf{u}_0)$. La forma en que T cambia un volumen estará reflejada en la forma en que $d_{u_0}T$ lo hace. En efecto, la traslación de un subconjunto por el vector $T(\mathbf{u}_0)$ no altera su volumen y como la diferencial $d_{u_0}T$ es una función lineal, cambia el volumen en una manera muy simple. En efecto, bajo la acción de una transformación lineal, los volúmenes quedan multiplicados por un factor constante que es el valor absoluto del determinante de la transformación. Por ejemplo, supongamos que T es una transformación lineal y tomemos como f la función constante 1. Entonces el teorema del cambio de variable (Teorema 3.1) implica el siguiente teorema.

3.2 Teorema

Si T es una transformación lineal de $\mathcal{R}^n$ en $\mathcal{R}^n$ con matriz A, entonces T multiplica volúmenes por el factor $|\det A|$.

Demostración. Haciendo $J = \det T'$ en el teorema del cambio de variable, tenemos que

$$V(T(R)) = \int_{T(R)} dV = \int_R |J|\,dV = |J|\,V(R.)$$

El último paso es válido, ya que para una transformación lineal con matriz A, el determinante jacobiano es constante igual a $\det A$. Entonces, $|J| = |\det A|$, el cual puede ser factorizado en la integral.

Ejemplo 5. La transformación de $\mathcal{U}$ en $\mathcal{R}$ definida por $x = 3u$ es lineal. Por tanto, ella es su diferencial (véase Cap. 3, Sec. 8, Ejercicio 9) y tiene un determinante jacobiano $J = 3$. De la Fig. 21 es claro que bajo esta transformación los largos quedan multiplicados por 3.

Ejemplo 6. La transformación T de $\mathcal{U}^2$ en $\mathcal{R}^2$ definida por

$$T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 \\ u + v \end{pmatrix}$$

tiene como diferencial en $\mathbf{u}_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ la transformación lineal

$$(d_{u_0}T)\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2u_0 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Cerca de

$$\mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

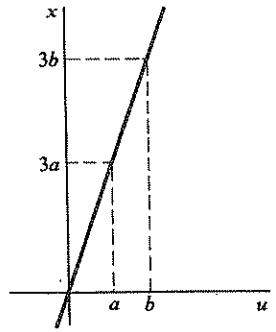


Figura 21

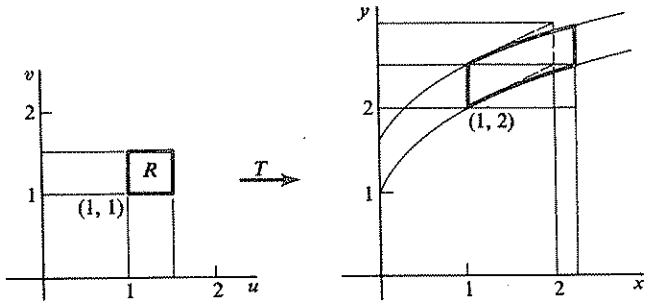


Figura 22

la función T es aproximada por la transformación afín

$$A\begin{pmatrix}u\\v\end{pmatrix} = T\begin{pmatrix}1\\1\end{pmatrix} + T'(\mathbf{u}_0)\begin{pmatrix}u-1\\v-1\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}2 & 0\\1 & 1\end{pmatrix}\begin{pmatrix}u-1\\v-1\end{pmatrix} = \begin{pmatrix}2u-1\\u+v\end{pmatrix}.$$

El cuadrado R en el plano uv de la Fig. 22 es transformado por T en la figura curva de la derecha. La transformación afín A transforma R en el paralelogramo indicado con líneas punteadas. Nótese que el área del paralelogramo es aproximadamente igual al área de la figura curva. Es fácil comprobar que el área del paralelogramo es $\frac{1}{2}$. Lo importante es que la aproximación afín de T duplica el área del cuadrado mientras que T aproximadamente duplica esa área. El factor de dilatación está dado por el determinante jacobiano de T en $\mathbf{u}_0 = \begin{pmatrix}1\\1\end{pmatrix}$. En efecto,

$$\det T' = \begin{pmatrix}2 & 0\\1 & 1\end{pmatrix} = 2.$$

Para calcular el área exacta de la imagen $T(R)$, usamos el teorema del cambio

de variables. Como u es positivo en R, la transformación T es inyectiva en R. La función inversa está dada por

$$T^{-1}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\sqrt{x}\\y-\sqrt{x}\end{pmatrix};$$

de modo que T es inyectiva. Además, su determinante jacobiano $J = 2u$ es positivo en R. Por tanto,

$$A(T(R)) = \int_{T(R)} dA = \int_R |J|\, dA$$
$$= \int_1^{3/2} dv \int_1^{3/2} 2u\, du = \tfrac{1}{2}u^2\Big]_1^{3/2} = \frac{5}{8}.$$

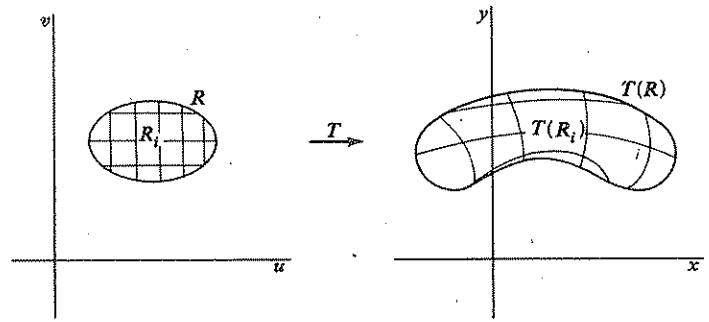


Figura 23

Para comprender el teorema del cambio de variables, sea T una transformación continuamente diferenciable y consideremos las regiones R y $T(R)$. La Fig. 23 ilustra un ejemplo en dos dimensiones. Descomponemos R en regiones R_i mediante rectas coordenadas. Si denotamos por $\approx$ la aproximación, tenemos que

$$\int_{T(R)} f\, dV = \sum_i \int_{T(R_i)} f\, dV \approx \sum_i f_i V(T(R_i)), \tag{2}$$

donde el número f_i es un valor que toma la función f en $T(R_i)$. Suponemos que $f_i V(TR_i)$ es una aproximación razonable de $\int_{T(R_i)} f dV$. Luego aproximamos $V(T(R_i))$ por $|J_i|V(R_i)$, donde J_i es un valor del determinante jacobiano de T en R_i. Llegamos así a la aproximación

$$\sum_i f_i V(T(R_i)) \approx \sum_i f_i\, |J_i|\, V(R_i). \tag{3}$$

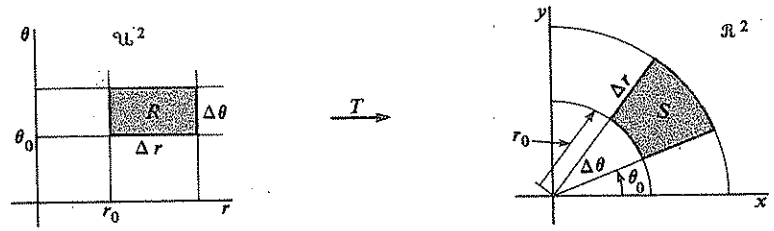


Figura 24

Ahora bien, el número f_i es también un valor de $f \circ T$ en R_i que podemos denotar por $(f \circ T)_i$ y de las Ecs. (2) y (3) obtenemos que

$$\int_{T(R)} f\,dV \approx \sum_i (f \circ T)_i\,|J_i|\,V(R_i).$$

Por último, esta última suma puede ser usada para aproximar $\int_R (f \circ T)\,|J|\,dV$.

Es difícil hacer preciso este argumento, por ello la demostración en el Apéndice sigue otras líneas.

Las consideraciones anteriores muestran que el determinante jacobiano puede ser interpretado como un factor aproximado de dilatación local del volumen. Si consideramos una transformación inyectiva continuamente diferenciable como un cambio de coordenadas, llegamos a una interpretación ligeramente diferente de J.

Ejemplo 7. Las coordenadas polares en $\mathfrak{R}^2$ acotan regiones como S de la Fig. 24. Como el determinante jacobiano de la transformación polar de coordenadas

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\mathrm{sen}\,\theta \end{pmatrix} = T\begin{pmatrix} r \\ \theta \end{pmatrix}$$

es

$$J = \begin{vmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{vmatrix} = r,$$

esperamos que $r_0\,\Delta r\,\Delta\theta$ es una aproximación al área sombreada $S = T(R)$ en el plano xy. Un cálculo exacto del área de S, usando el teorema del cambio de variables da

$$\begin{aligned} \int_S dA = \int_R r\,dA &= \int_{r_0}^{r_0+\Delta r} r\,dr \int_{\theta_0}^{\theta_0+\Delta\theta} d\theta \\ &= \left(\tfrac{1}{2}(r_0+\Delta r)^2 - \tfrac{1}{2}r_0^2\right)\Delta\theta \\ &= r_0\,\Delta r\,\Delta\theta + \tfrac{1}{2}(\Delta r)^2\,\Delta\theta \\ &\approx r_0\,\Delta r\,\Delta\theta \qquad \text{(para pequeño } \Delta r, \Delta\theta\text{)}. \end{aligned}$$

Por tanto, el significado de J en este caso es que $J\,\Delta r\,\Delta\theta$ es una aproximación del área de la región acotada por curvas coordenadas tales que Δr y $\Delta\theta$ son las diferencias entre pares de valores de r y θ.

La expresión $r\,\Delta r\,\Delta\theta$ se llama el **elemento de área** en coordenadas polares. Más generalmente, si J es el determinante de un cambio de coordenadas, entonces $|J|\,\Delta V$ se llama su **elemento de volumen**.

Ejemplo 8. Se introducen coordenadas esféricas en $\mathfrak{R}^3$ mediante la transformación

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\,\mathrm{sen}\,\phi\cos\theta \\ r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta \\ r\cos\phi \end{pmatrix}.$$

Esta misma transformación ha sido considerada en el Ejemplo 3 con otra notación. El determinante jacobiano es $J = r^2\,\mathrm{sen}\,\phi$. Esto sugiere que una aproximación para el volumen del «cubo» C en coordenadas esféricas de la Fig. 25, es $r^2\,\mathrm{sen}\,\phi\,\Delta r\,\Delta\phi\,\Delta\theta$.

La bola con centro en el origen y radio 1 en $\mathfrak{R}^3$ está definida por $x^2 + y^2 + z^2 \le 1$ y la denotamos por B_{xyz}. Esta misma bola está definida en coordenadas esféricas por las desigualdades

$$0 \le r \le 1, \qquad 0 \le \phi \le \pi, \qquad 0 \le \theta \le 2\pi,$$

y la denotamos por $B_{r\phi\theta}$. Usando el teorema del cambio de variables, podemos calcular el volumen de la bola:

$$\begin{aligned} V(B_{xyz}) = \int_{B_{xyz}} dx\,dy\,dz &= \int_{B_{r\phi\theta}} r^2\,\mathrm{sen}\,\phi\,dr\,d\phi\,d\theta \\ &= \int_0^1 r^2\,dr \int_0^\pi \mathrm{sen}\,\phi\,d\phi \int_0^{2\pi} d\theta \\ &= \tfrac{1}{3}(1+1)2\pi = \frac{4\pi}{3}. \end{aligned}$$

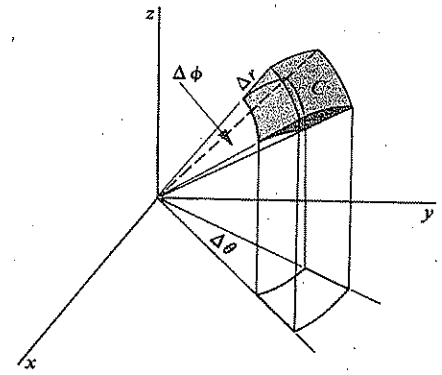


Figura 25

Nótese que (tal como en el Ejemplo 3) las condiciones (a) y (b) del Teorema 3.1 fallan en $B_{r\phi\theta}$. Sin embargo, salvo en conjuntos de volumen cero, el determinante jacobiano $J = r^2 \operatorname{sen} \phi$ es positivo y la transformación de coordenadas es inyectiva. Por tanto, la aplicación del teorema del cambio de variable es legítima.

EJERCICIOS

1. Sea

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 - v^2 \\ 2uv \end{pmatrix}.$$

(a) Dibuje la imagen por T del cuadrado en $\mathfrak{U}^2$ con vértices $(1, 1)$, $(1, \frac{3}{2})$, $(\frac{3}{2}, 1)$, $(\frac{3}{2}, \frac{3}{2})$.

(b) Dibuje la imagen por $T'\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ del cuadrado de la parte (a).

(c) Dibuje el trasladado de la imagen encontrada en la parte (b) por el vector

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} - T'\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Verifique que esta es la imagen del cuadrado por la aproximación afín de T en $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(d) Encuentre el área de la región dibujada en (c). [*Resp.* 2.]

(e) Encuentre el área de la región dibujada en (a). [*Resp.* $\frac{19}{6}$.]

2. Sea

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \end{pmatrix}.$$

(a) Dibuje la imagen por T del cuadrado S con vértices en $(0, 0)$, $(0, \pi/2)$, $(\pi/2, 0)$ y $(\pi/2, \pi/2)$.

(b) Dibuje la imagen por $T'\begin{pmatrix} \pi/2 \\ 0 \end{pmatrix}$ del cuadrado S. ¿Cuál es el área de esta imagen?

(c) Dibuje la imagen de S por la aproximación afín de T en $(\pi/4, \pi/4)$. ¿Cuál es su área?

(d) ¿Cuál es el área de la región dibujada en (a)?

3. Sea

$$T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \end{pmatrix}.$$

Demuestre que $T'\begin{pmatrix} u \\ v \end{pmatrix}$ transforma un rectángulo de área A en una región de área uA.

4. Calcule el área de la imagen del rectángulo en el plano uv con vértices $(0, 0)$, $(0, 1)$, $(2, 0)$ y $(2, 1)$ por la transformación

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$ [*Resp*: 8.]

5. Considere la transformación T definida por

$$\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^2 - v^2 \\ 2uv \end{pmatrix}.$$

Sea R_{uv} la parte del disco unitario que yace en el primer cuadrante; es decir, $u^2 + v^2 \leq 1$, $u \geq 0$, $v \geq 0$.

(a) Dibuje la región imagen $R_{xy} = T(R_{uv})$.

(b) Calcule $\displaystyle\int_{R_{xy}} \frac{dx\,dy}{\sqrt{x^2 + y^2}}$. [*Resp.* π.]

6. Considere la transformación del plano xy en el plano uv definida por $x = u + v$, $y = u^2 - v$. Sea R_{uv} la región acotada por (1) el eje u, (2) el eje v, (3) la recta $u + v = 2$.

(a) Encuentre y dibuje la región imagen R_{xy}.

(b) Calcule la integral

$$\int_{Rxy} \frac{dx\,dy}{\sqrt{1 + 4x + 4y}}.$$ [*Resp*: 2.]

7. Considere la transformación del plano uv en el plano xy definida por

$$x = u, \qquad y = v(1 + u^2),$$

y sea R_{uv} la región rectangular definida por $0 \leq u \leq 3$ y $0 \leq v \leq 2$.

(a) Encuentre y dibuje la región imagen R_{xy}.

(b) Encuentre $\dfrac{\partial(x, y)}{\partial(u, v)}$.

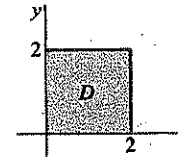
(c) Transforme la integral $\int_{R_{xy}} x\,dx\,dy$ en una integral en R_{uv} y calcule las dos integrales. [*Resp.* $\frac{99}{2}$.]

8. La transformación $u = x^2 - y^2$, $v = 2xy$ transforma la región D (véase dibujo) en una región R en el plano uv y es inyectiva en D.

(a) Encuentre R.

(b) Calcule $\int_R 1\,du\,dv$ primero integrando directamente en R y después usando las fórmulas de transformación para integrar en D. [*Resp.* $\frac{128}{3}$.]

(c) Calcule $\int_R v\,du\,dv$ directamente y usando el teorema del cambio de variables. [*Resp.* 128.]



9. Considere la transformación del plano xy en el plano uv definida por

$$u = x$$
$$v = y(1 + 2x).$$

(a) ¿Qué le sucede a las rectas horizontales del plano xy?

(b) Si D es la región rectangular

$$0 \le x \le 3$$
$$1 \le y \le 3,$$

encuentre la región R, imagen de D.

(c) Calcule

$$\int_R du\,dv, \qquad \int_R v\,dv\,du, \qquad \text{y} \qquad \int_R u\,dv\,du$$

por integración directa y después reduciendo a integrales en D. [*Resp.* 24, 228, 45.]

10. Calcule el área acotada por las curvas en coordenadas polares $\theta = 0$, $\theta = \pi/4$ y $r = \theta^2$. [*Resp.* $\pi^5/(2^{10} \cdot 10)$.]

11. Calcule el área acotada por la lemniscata $(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$ pasando a coordenadas polares. [*Resp.* $2a^2$.]

12. Calcule el volumen del elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \le 1.$$

[Use la transformación $(x, y, z) = (au, bv, cw)$ para transformar la esfera $u^2 + v^2 + w^2 \le 1$ en el elipsoide. Suponga conocido el volumen de la esfera.]

13. Evalúe la integral de $f(x, y, z) = a$ en el hemisferio $x^2 + y^2 + z^2 \le 1$, $x \ge 0$, pasando a coordenadas esféricas. [*Resp.* $\frac{2}{3}\pi a$.]

14. (a) Calcule el jacobiano de la transformación de coordenadas cilíndricas

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\,\text{sen}\,\theta \\ z \end{pmatrix}.$$

(b) Use coordenadas cilíndricas para calcular

$$\int_{\substack{0\le z\le 1 \\ x^2+y^2\le 1,}} x^2\,dx\,dy\,dz.$$

15. Demuestre que la transformación

$$x_1 = u_1$$
$$x_2 = u_1 + u_2$$
$$x_3 = u_1 + u_2 + u_3$$
$$\vdots$$
$$x_n = u_1 + u_2 + \ldots + u_n$$

no cambia volúmenes.

16. Si cortamos un sólido de revolución R en conchas cilíndricas delgadas con el mismo eje que el eje de revolución, entonces obtenemos intuitivamente la fórmula

$$\int 2\pi r h(r)\,dr.$$

para el volumen del sólido. En esta fórmula, $h(r)$ es el grosor del sólido a una distancia r del eje, medida a lo largo de una recta paralela al eje. Demuestre que si introducimos coordenadas cilíndricas en la integral $\int_R dx\,dy\,dz$, obtenemos la misma fórmula.

17. (a) Suponga que la bola B de radio a tiene una densidad ρ en cada uno de sus puntos igual a la distancia del punto a un diámetro fijo. Encuentre la masa total de la bola.
[*Indicación*: Calcule la integral $\int_B \rho\,dV$ usando coordenadas esféricas.]

(b) Suponga que un cilindro de alto h y radio a tiene una densidad ρ en cada punto igual a la distancia del punto al eje del cilindro. Calcule la masa total del cilindro.

4. Integrales impropias

La definición de la integral puede ser extendida a funciones no acotadas y que no son necesariamente cero fuera de un conjunto acotado. Primero consideraremos algunos ejemplos.

Ejemplo 1. La función $f(x, y) = 1/x^2y^2$, definida para $x \ge 1$ e $y \ge 1$, tiene el gráfico de la Fig. 26. Si B es el conjunto de los puntos (x, y) para los cuales $x \ge 1$ e $y \ge 1$, es natural definir $\int_B f\,dA$ de manera que pueda ser llamada el volumen bajo el gráfico de f. Podemos aproximar este volumen calculando el volumen sobre subrectángulos acotados de B. Específicamente, sea B_N el rectángulo con esquinas $(1, 1)$ y (N, N) y aristas paralelas a las de B. Para $N > 1$, tenemos

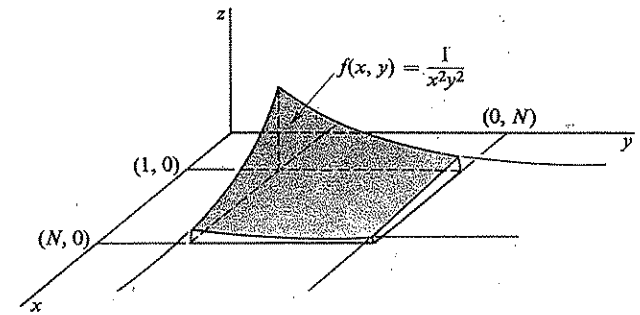


Figura 26

$$\int_{B_N} f\,dA = \int_1^N dx \int_1^N \frac{1}{x^2y^2}\,dy$$
$$= \left(\int_1^N \frac{dx}{x^2}\right)^2 = \left(1 - \frac{1}{N}\right)^2.$$

Cuando N tiende a infinito, los rectángulos B_N cubren eventualmente todo punto de B y las regiones B_N llevan toda la región bajo el gráfico de f. Entonces, por definición,

$$\int_B f\,dA = \lim_{N\to\infty} \int_{B_N} f\,dA = 1.$$

Ejemplo 2. Sea B el disco $x^2 + y^2 \leq 1$ en $\mathfrak{R}^2$ y supongamos que

$$f(x, y) = -\log(x^2 + y^2), \qquad 0 < x^2 + y^2 \leq 1.$$

La Fig. 27 muestra el gráfico de f. Como f no es acotada cerca de $(0, 0)$, sacamos de B un disco de centro $(0, 0)$ y radio ϵ. Llamamos B_ϵ la parte restante de B. Usando coordenadas polares, tenemos que

$$\begin{aligned}\int_{B_\epsilon} -\log(x^2 + y^2)\,dx\,dy &= \int_0^{2\pi} d\theta \int_\epsilon^1 -(\log r^2) r\,dr \\ &= -2\pi[r^2 \log r - \tfrac{1}{2}r^2]_\epsilon^1 \\ &= \pi + 2\pi\epsilon^2 \log\epsilon - \pi\epsilon^2.\end{aligned}$$

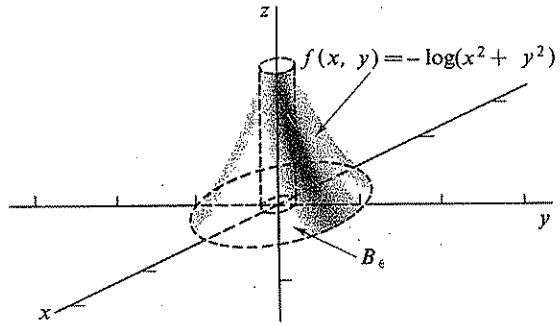


Figura 27

Como $\lim_{\epsilon\to 0} (2\pi\epsilon^2 \log\epsilon - \pi\epsilon^2) = 0$, tenemos, por definición,

$$\int_B -\log(x^2 + y^2)\,dx\,dy = \lim_{\epsilon\to 0} \int_{B_\epsilon} -\log(x^2 + y^2)\,dx\,dy = \pi.$$

En el Ejemplo 2, la función $-\log(x^2 + y^2)$ no es acotada en el disco $x^2 + y^2 \leq 1$ solo alrededor del punto $(0, 0)$. Evidentemente, es importante encontrar todos los puntos de este tipo al tratar de integrar una función no acotada. En general, decimos que un punto x es un **punto de discontinuidad infinita** de f si en toda vecindad de x, $|f|$ toma valores arbitrariamente grandes.

Ejemplo 3. Consideremos la función $g(x, y, z) = (x^2 + y^2 + z^2 - 1)^{-1/2}$ para (x, y, z) tal que $1 < x^2 + y^2 + z^2 \leq 2$. El dominio de g es la región entre

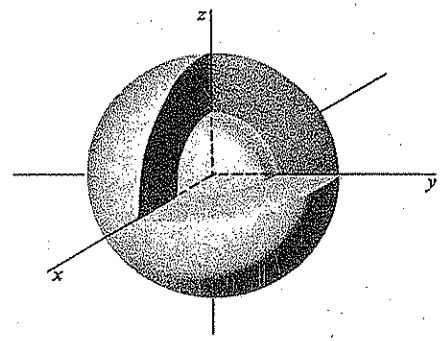


Figura 28

las dos esferas concéntricas mostradas en la Fig. 28. Todos los puntos de la esfera $x^2 + y^2 + z^2 = 1$ son puntos de discontinuidad infinita de g. Para definir la integral de g, aproximamos su dominio por las conchas B_ϵ determinadas por $1 + \epsilon \leq x^2 + y^2 + z^2 \leq 2$. Estas conchas tienen la propiedad de que cubren totalmente el dominio de g cuando ϵ tiende a cero, a pesar de que ninguna de ellas contiene un punto de discontinuidad infinita. Introduciendo coordenadas esféricas, tenemos que

$$\begin{aligned}\int_{B_\epsilon} (x^2 + y^2 + z^2 - 1)^{-1/2}\,dx\,dy\,dz &= \int_{1+\epsilon}^2 dr \int_0^{2\pi} d\theta \int_0^\pi (r^2 - 1)^{-1/2} r^2 \operatorname{sen}\phi\,d\phi \\ &= 4\pi \int_{1+\epsilon}^2 (r^2 - 1)^{-1/2} r^2\,dr \\ &= 4\pi[\tfrac{1}{2}r\sqrt{r^2 - 1} + \tfrac{1}{2}\log(r + \sqrt{r^2 - 1})]_{1+\epsilon}^2.\end{aligned}$$

De lo cual sigue inmediatamente que

$$\lim_{\epsilon\to 0} \int_{B_\epsilon} (x^2 + y^2 - z^2 + 1)^{1/2}\,dx\,dy\,dz = 4\pi\sqrt{3} + 2\pi\log(2 + \sqrt{3}).$$

Antes de resumir en una definición las ideas recién ilustradas, impondremos condiciones para el integrando f y para el conjunto B, en el cual se va a integrar.

1. Dea D el conjunto de los puntos de B en los cuales f no es continua. La parte de D contenida en un rectángulo acotado arbitrario debe estar contenida en un número finito de conjuntos lisos.

2. La parte de B contenida en un rectángulo acotado arbitrario debe tener una frontera formada por un número finito de conjuntos lisos.

Ambas condiciones son satisfechas para los tres ejemplos que hemos conside-

rado hasta ahora y supondremos que ellos son válidos en todos los casos considerados en el resto de esta sección.

En los ejemplos hemos visto que la integral $\int_B f\,dV$ puede ser definida en algunos casos en que f o B no son acotados. La definición general de la integral será dada de modo que ambos fenómenos puedan ocurrir al mismo tiempo. Procedemos como sigue: Se dirá que una familia creciente $\{B_N\}$ de subconjuntos de B **converge a** B si todo subconjunto acotado de B, en el cual f es acotada, está contenido en alguno de los subconjuntos B_N. Nótese que esta definición de convergencia depende no solo de B, sino que también de f. El índice N puede ser elegido de alguna forma conveniente, por ejemplo, puede tender a ∞, ya sea continuamente ya sea tomando valores enteros, o puede tender a un número finito. En el resto de esta sección supondremos que en toda familia creciente $\{B_N\}$ que converge a B, todos los subconjuntos B_N satisfacen la condición 2.

La **integral de** f **en** B es, por definición,

$$\lim_N \int_{B_N} f\,dV = \int_B f\,dV,$$

siempre que el límite sea finito y no dependa de la familia de conjuntos acotados B_N que converge a B. Se supone que los B_N se eligen de modo que las integrales ordinarias de Riemann $\int_{B_N} f\,dV$ (como han sido definidas en la Sec. 2) existen. La integral obtenida de esta manera se llama la **integral impropia de Riemann** cuando es necesario distinguirla de la integral de Riemann de una función acotada en un conjunto acotado.

A pesar de que el requisito de que el valor de la integral sea independiente de la familia convergente de conjuntos usada para definirla, es natural, más adelante veremos que a veces es interesante descartarlo. Sin embargo, el teorema siguiente muestra que, para funciones positivas, el límite de $\int_{B_N} f\,dV$ es siempre independiente de la familia de conjuntos.

4.1 Teorema

Sea f no negativa en B y supongamos que

$$\lim_N \int_{B_N} f\,dV$$

es finito para una cierta familia creciente de conjuntos B_N que converge a B. Entonces $\int_B f\,dV$ está definida y su valor es igual a

$$\lim_N \int_{C_N} f\,dV,$$

donde $\{C_N\}$ es cualquier familia que converge a B.

Demostración. Como f es acotada en cada B_N, para cada N existe un índice K tal que

$$B_N \subset C_K.$$

De la misma manera, existe un índice M que depende de K tal que

$$C_K \subset B_M.$$

Entonces, como f es no negativa,

$$\int_{B_N} f\,dV \leq \int_{C_K} f\,dV \leq \int_{B_M} f\,dV.$$

Además,

$$\int_{C_N} f\,dV \leq \lim_N \int_{B_N} f\,dV$$

para todo N. Como $\int_{C_N} f\,dV$ es creciente y acotada superiormente,

$$\lim_N \int_{C_N} f\,dV$$

existe. La doble desigualdad muestra que

$$\lim_N \int_{B_N} f\,dV = \lim_N \int_{C_N} f\,dV.$$

Esto completa la demostración.

Ejemplo 4. Sea f definida por $f(x, y) = y^{-1/2}e^{-x}$ en la cinta infinita S en $\mathcal{R}^2$ mostrada en la Fig. 29. Es claro que f tiene una discontinuidad infinita

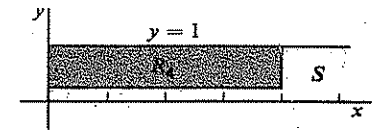


Figura 29

en cada punto del eje x positivo. Definimos R_N como el rectángulo en S acotado por las rectas $x = N$ e $y = 1/N$, para $N > 1$. Cuando N tiende a infinito, R_N converge a S. Tenemos que

$$\begin{aligned}\int_{R_N} f\,dA &= \int_0^N dx \int_{1/N}^1 y^{-1/2}e^{-x}\,dy \\ &= \left(\int_0^N e^{-x}\,dx\right)\left(\int_{1/N}^1 y^{-1/2}\,dy\right) \\ &= (1 - e^{-N})\left(2 - \frac{2}{\sqrt{N}}\right).\end{aligned}$$

Entonces

$$\int_S y^{-1/2}e^{-x}\,dx\,dy = \lim_{N\to\infty}(1 - e^{-N})\left(2 - \frac{2}{\sqrt{N}}\right) = 2.$$

Ejemplo 5. La integral de $1/x^a$ en el eje x positivo, denotada por $\int_0^\infty x^{-a}\,dx$, no existe para todo a. Consideremos, para todo $N > 0$,

$$\int_{1/N}^{N} x^{-a}\,dx = \begin{cases} \dfrac{N^{1-a} - 1/N^{1-a}}{1-a}, & a \neq 1, \\ 2\log N, & a = 1. \end{cases}$$

Cuando N tiende a infinito, en cada caso el límite es infinito. Sin embargo, es fácil verificar que

$$\int_1^\infty x^{-a}\,dx = \frac{1}{a-1}, \quad \text{para } a > 1,$$

y

$$\int_0^1 x^{-a}\,dx = \frac{1}{1-a}, \quad \text{para } a < 1.$$

La integral

$$\int_{-1}^{1} \frac{1}{x}\,dx$$

no existe si exigimos que su valor sea independiente del proceso de límite por el cual se ha calculado. En efecto, si integramos primero en los intervalos $[-1, -\delta]$ y $[\epsilon, 1]$ con $0 < \delta < 1$ y $0 < \epsilon < 1$, obtenemos

$$\int_{-1}^{-\delta} \frac{1}{x}\,dx + \int_{\epsilon}^{1} \frac{1}{x}\,dx = \log|x|\Big]_{-1}^{-\delta} + \log x\Big]_{\epsilon}^{1}$$

$$= \log\frac{\delta}{\epsilon}.$$

Cuando ϵ y δ tienden a cero, $\log(\delta/\epsilon)$ se puede hacer tender a cualquier número controlando el límite de δ/ϵ. En particular, si hacemos $\epsilon = \delta$, el límite es cero.

Para una función f con un gráfico simétrico con respecto de un punto $\mathbf{x}_0$, a veces tiene sentido definir $\int_B f\,dV$ como un límite usando conjuntos B_N que también son simétricos con respecto de $\mathbf{x}_0$. Esto es lo que hemos hecho en el ejemplo anterior y en general hablamos de calcular un **valor principal** (v.p.) de la integral. Para la integral en la última parte del Ejemplo 5 escribimos

$$\text{v.p.}\int_{-1}^{1} \frac{dx}{x} = 0.$$

Ejemplo 6. Sean (r, θ) las coordenadas polares en $\mathfrak{R}^2$ y sea $f(r, \theta) = (\operatorname{sen}\theta)/r^2$ en el disco D de radio 1 y centro en el origen. Es claro que f tiene una discontinuidad infinita en el origen, ya que, por ejemplo, a lo largo de la recta $\theta = \pi/2$, f tiende a ∞ y a lo largo de la recta $\theta = 3\pi/2$, f tiende a $-\infty$. (La Fig. 30 muestra el gráfico de f.) Sin embargo, $\int_D f\,dA$ no existe en el sentido de la integral impropia ordinaria, puesto que el límite que se obtiene usando una sucesión de regiones en D depende de la forma en que se equilibran los

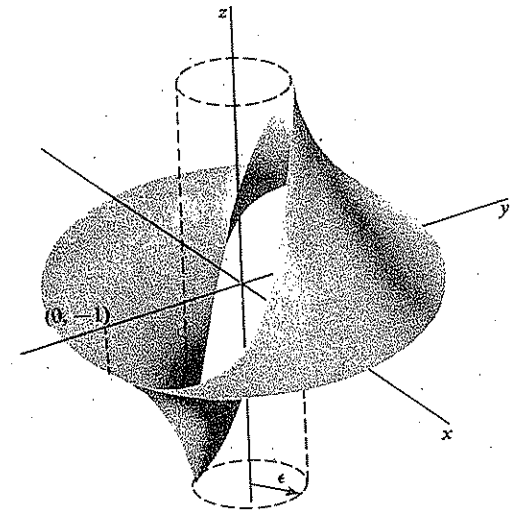


Figura 30

valores positivos y negativos de f. Se puede determinar un valor principal tomando el límite con respecto a una familia de regiones anulares. Sea D_ϵ el anillo $\epsilon \leq r \leq 1$. Entonces

$$\text{v.p.}\int_D f\,dV = \lim_{\epsilon\to 0}\int_{D_\epsilon} \frac{\operatorname{sen}\theta}{r^2}\,r\,dr\,d\theta$$

$$= \lim_{\epsilon\to 0}\int_0^{2\pi} \operatorname{sen}\theta\,d\theta \int_\epsilon^1 \frac{dr}{r} = 0.$$

El teorema siguiente es un criterio conveniente para la existencia de una integral impropia.

4.2 Teorema

Supongamos que f y g tienen los mismos puntos de discontinuidad infinita. Si $|f| \leq g$ y $\int_B g\,dV$ existe, entonces $\int_B f\,dV$ también existe.

Demostración. Sea $\{B_N\}$ una familia creciente de conjuntos que converge a B. Como $f + |f| \leq 2|f| \leq 2g$, tenemos que

$$\int_{B_N} (f + |f|)\,dV \leq 2\int_{B_N} g\,dV \leq 2\int_B g\,dV.$$

Entonces, ya que $f + |f| \geq 0$, el valor de la integral $\int_{B_N} (f + |f|)\,dV$ aumenta a medida que B_N lo hace, tenemos que

$$\lim_N \int_{B_N} (f + |f|)\, dV = l_1 \leq 2\int_B g\, dV.$$

Análogamente,

$$\lim_N \int_{B_N} |f|\, dV = l_2 \leq \int_B g\, dV.$$

Por último,

$$\lim_N \int_{B_N} f\, dV = \lim_N \left(\int_{B_N} (f + |f|)\, dV - \int_{B_N} |f|\, dV \right)$$
$$= \lim_N \int_{B_N} (f + |f|)\, dV - \lim_N \int_{B_N} |f|\, dV = l_1 - l_2.$$

Como la familia B_N es arbitraria, $\int_B f\, dV$ está definida.

Ejemplo 7. Sea B el disco $x^2 + y^2 \leq 1$ en $\mathcal{R}^2$ y sea f definida por

$$f(x, y) = \begin{cases} (x^2 + y^2)^{-1/2}, & \text{para } x \geq 0 \quad \text{y} \quad x^2 + y^2 > 0. \\ (x^2 + y^2)^{1/2}, & \text{para } x < 0. \end{cases}$$

Como $\int_B (x^2 + y^2)^{-1/2}\, dx\, dy$ existe, y $|f(x, y)| \leq (x^2 + y^2)^{-1/2}$ para $0 < x^2 + y^2 < 1$, del Teorema 4.2 se sigue que $\int_B f\, dx\, dy$ existe. Se deja como ejercicio el cálculo de esta integral.

Ejemplo 8. Sea $j(x) = (-1)^{n-1}/n$ para $n - 1 \leq x \leq n$ y $n = 1, 2, 3, \ldots$ El gráfico de j está representado en la Fig. 31 hasta $x = 4$. Entonces

$$\lim_{n\to\infty} \int_0^n j(x)\, dx = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2,$$

y podemos escribir

$$\int_0^\infty j(x)\, dx = \log 2,$$

si se entiende que el pasaje al límite ha sido efectuado en esta forma especial. Este ejemplo comparte con los ejemplos sobre el valor principal la propiedad de que el valor asignado a la integral de la elección de una sucesión de regiones son respecto a la cual se ha tomado el límite. Una integral de este tipo se llama **condicionalmente convergente.** Otro ejemplo está dado en el Ejercicio 8.

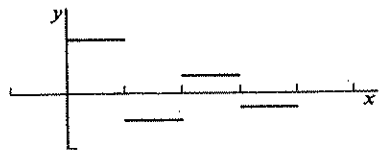


Figura 31

EJERCICIOS

1. En cada parte, determine si la integral está definida o no. Si ella está definida, calcule su valor.

(a) $\int_0^\infty \frac{dx}{x^2 + 1}$. [*Resp*: $\pi/2$.]

(b) $\int_{-\infty}^\infty \frac{dx}{x^2 - 1}$.

(c) $\int_0^1 \frac{dx}{\sqrt{1 - x^2}}$.

(d) $\int_{x^2+y^2\leq 1} \frac{dx\, dy}{\sqrt{x^2 + y^2}}$.

(e) $\int_R \frac{(x - y)\, dx\, dy}{x^2 + y^2}$, donde R es el rectángulo $\max(|x|, |y|) \leq 1$.

(f) $\int_{x^2+y^2+z^2\geq 1} \frac{dx\, dy\, dz}{(x^2 + y^2 + z^2)^2}$.

(g) $\int_{x^2+y^2+z^2\geq 1} \frac{dx\, dy\, dz}{xyz}$.

(h) $\int_C e^{-x-y-z}\, dx\, dy\, dz$, donde C es la columna infinita $\max(|x|, |y|) \leq 1, \quad z \geq 0$.

2. Demuestre que

(a) $$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1}\, dx = (n - 1)!$$

donde $n > 1$ es un entero.

(b) Exprese $\int_T e^{-x}(x - y)^{-1/2}\, dx\, dy$ en términos de Γ, donde T es la región $x \geq y \geq 0$. [*Resp*. $2\Gamma(\frac{3}{2})$.]

3. Sea B la bola $|\mathbf{x}| \leq 1$ en $\mathcal{R}^n$. ¿Para cuáles valores de a existe $\int_B \frac{dV}{|\mathbf{x}|^a}$?

4. Calcule (a) v.p. $\int_{-\infty}^\infty \frac{x\, dx}{x^2 + 1}$. (b) v.p. $\int_{-1}^1 \frac{x\, dx}{2x - 1}$.

5. Calcule los valores de la función $g(y) = \int_{-1}^1 \frac{x\, dx}{x - y}$, tomando un valor principal de la integral cuando sea necesario.

6. Calcule la integral de la función f del Ejemplo 7 y calcule $\int_B (x^2 + y^2)^{-1/2}\, dx\, dy$. [*Resp*. $\frac{4}{3}\pi$, 2π.]

7. Demuestre que la integral del Ejemplo 8 depende de la sucesión de conjuntos usada para calcular el límite. [*Sugerencia*: Tome cada B_N como una unión disconexa de intervalos.]

8. Sea $f(x, y) = \operatorname{sen}(x^2 + y^2)$ en el cuadrante Q definido por $x \geq 0, y \geq 0$. Demues-

tre que $\int_Q f\,dA$ converge condicionalmente. [*Sugerencia*: Para obtener un límite integre sobre cuadrados crecientes. Luego integre sobre cuartos de disco.]

9. ¿En qué sentido existe cada una de las integrales siguientes? Las posibilidades son: integral de Riemann ordinaria, integral impropia, integral condicionalmente convergente y ninguna de las anteriores.

(a) $\int_\pi^\infty \frac{\operatorname{sen} x}{x^2}\,dx$. (c) $\int_0^\infty \operatorname{sen} x\,dx$.

(b) $\int_0^\infty \frac{\operatorname{sen} x}{x}\,dx$. (d) $\int_0^1 \operatorname{sen}\frac{1}{x}\,dx$.

10. La integral $\int_0^\infty f(x)\,dx$ se dice es **sumable en el sentido de Abel** con valor k si

$$\lim_{\epsilon\to 0+}\int_0^\infty e^{-\epsilon x}f(x)\,dx = k.$$

Encuentre este valor para (a) y (b). (a) $\int_0^\infty \operatorname{sen} x\,dx$, (b) $\int_0^\infty \cos x\,dx$.

11. (a) Calcule $\int_{\mathscr{R}^2} e^{-x^2-y^2}\,dx\,dy$. (Use coordenadas polares.)

(b) Use el resultado de la parte (a) para calcular $\int_{-\infty}^{\infty} e^{-x^2}\,dx$.

(c) Calcule $\int_{\mathscr{R}^n} \exp(-x_1^2 - \ldots - x_n^2)\,dV$.

12. (a) Demuestre que el área acotada por el gráfico de $y = 1/x$, el eje x y la recta $x = 1$ es infinita.

(b) Calcule el volumen barrido por la región descrita en la parte (a) al rotar alrededor del eje x.

13. Sea f positiva y no acotada en un conjunto no acotado B en $\mathscr{R}^2$. Considere la región C entre el gráfico de f y B y demuestre que si

$$\int_B f\,dA \quad \text{y} \quad \int_C dV$$

existen, ellas son iguales.

14. Demuestre que si $\int_B |f|\,dV$ existe, entonces también existe $\int_B f\,dV$. Si no están satisfechas las condiciones (1) y (2), esto es falso. Por ejemplo, B es el intervalo unitario $0 \le x \le 1$ y f la función

$$f(x) = \begin{cases} 1, & \text{si } x \text{ es racional.} \\ -1, & \text{si } x \text{ es irracional.} \end{cases}$$

15. Demuestre que si la integral de Riemann ordinaria $\int_B f\,dV$ existe, entonces ella también existe como integral impropia (dadas las condiciones (1) y (2)) y que ambas integrales son iguales.

16. Una función no negativa $\mathscr{R}^n \xrightarrow{p} \mathscr{R}$ integrable en una región R de $\mathscr{R}$ se llama una **densidad de probabilidad** si $\int_R p\,dV = 1$. La **media** de p se define como el vector

$$M[p] = \int_R \mathbf{x}p(\mathbf{x})\,dV,$$

y la **varianza** de p es el número real

$$\sigma^2[p] = \int_R |\mathbf{x} - M[p]|^2 p(\mathbf{x})\,dV.$$

Demuestre que cada una de las funciones siguientes es una densidad de probabilidad y calcule su media y varianza si ellas existen.

(a) $p(x, y) = (1/2\pi)e^{-(x^2+y^2)/2}$. [*Indicación*: Use coordenadas polares.]

(b) $p(x) = (\pi(1 + x^2))^{-1}$.

5. Estimación de integrales

En varios de los ejemplos de este capítulo la evaluación de las integrales se ha hecho usando el teorema fundamental del cálculo para llegar a una respuesta precisa en términos de funciones elementales. En la práctica, muy a menudo no es posible efectuar este tipo de cálculo y entonces solo podremos hacer una estimación de la integral. La desigualdad fundamental que se usa para hacer estimaciones es la del Teorema 2.8, que repetimos a continuación.

5.1. Si f y g son integrales en B y $f \le g$ en B, entonces

$$\int_B f\,dV \le \int_B g\,dV.$$

Ejemplo 1. La función f de una variable definida por

$$f(x) = \frac{\operatorname{sen} x}{(1 + x^2)^2}, \qquad 0 \le x \le 1$$

tiene una integral difícil de calcular. Comparándola con los gráficos de x y de sen x para $0 \le x \le 1$ se tiene que, como lo muestra la Fig. 32, $0 \le \operatorname{sen} x \le x$ para los valores indicados de x. Entonces, como $(1 + x^2)^2 > 0$,

$$0 \le \frac{\operatorname{sen} x}{(1 + x^2)^2} \le \frac{x}{(1 + x^2)}.$$

Por tanto, por 5.1,

$$0 \le \int_0^1 \frac{\operatorname{sen} x}{(1 + x^2)^2}\,dx \le \int_0^1 \frac{x\,dx}{(1 + x^2)^2}.$$

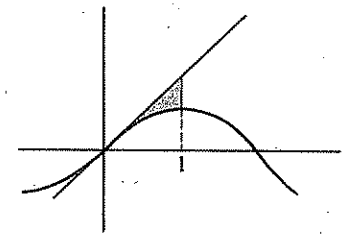


Figura 32

Pero esta última integral es fácil de calcular

$$\int_0^1 \frac{x\,dx}{(1+x^2)^2} = \left[\frac{-1}{2(1+x^2)}\right]_0^1 = -\tfrac{1}{4} + \tfrac{1}{2} = 0{,}25;$$

Por tanto,

$$0 \le \int_0^1 \frac{\operatorname{sen} x}{(1+x^2)^2}\,dx \le 0{,}25.$$

Ejemplo 2. Supongamos que queremos estimar la integral de $f(x, y) = \cos(x + y/2)$ en el cuadrado S: $0 \le x \le \frac{1}{2}$, $0 \le y \le \frac{1}{2}$.

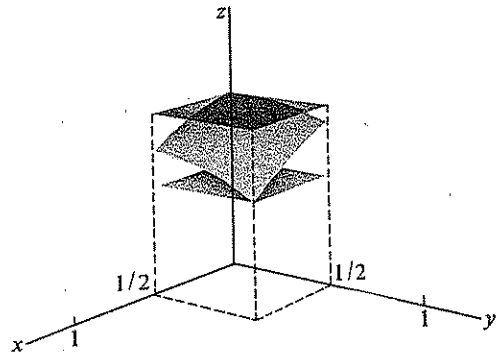


Figura 33

La Fig. 33 muestra el gráfico de f sobre los planos horizontales en $z = 1$ y $z = \cos\frac{3}{4} > 0{,}73$. Por tanto,

$$0{,}73 \le \cos\left(\frac{x+y}{2}\right) \le 1.$$

Por 5.1 tenemos que

$$\int_S 0{,}73\,dx\,dy \le \int_S \cos\left(\frac{x+y}{2}\right) dx\,dy \le \int_S 1\,dx\,dy.$$

La evaluación de la integral más pequeña y de la más grande es fácil, ya que ellas son los volúmenes de rectángulos sólidos de base S y alturas 0,73 y 1, respectivamente. Como el área de S es 0,25, se tiene que

$$0{,}1825 \le \int_S \cos(x + 2y)\,dx\,dy \le 0{,}25.$$

Las estimaciones de los dos ejemplos anteriores son bastante burdas, puesto que hemos remplazado el integrando f por funciones que son bastante distintas de f en una parte bastante grande del dominio de integración. Las estimaciones son mejores si se usan aproximaciones mejores de f. Sin embargo, las funciones aproximadas deben ser fáciles de integrar. Una manera de lograr esto es eligiendo como aproximación una función escalonada constante en cada subrectángulo de un reticulado rectangular en el dominio de integración. La función escalonada puede ser elegida de tal modo que su integral sea una suma de Riemann. En seguida se puede encontrar una sucesión de estas funciones escalonadas cuyas integrales converjan a la integral de f.

Ejemplo 3. La integral

$$\int_R (1 + x + y)^3\,dx\,dy,$$

donde R es el rectángulo $0 \le x \le 1$, $0 \le y \le \frac{1}{2}$ puede ser aproximada por sumas de Riemann. Subdividimos el rectángulo mediante un reticulado con vértices en los puntos $(j/N, k/2M)$ donde $j = 1, \ldots, N$, $k = 1, \ldots, M$. Al evaluar el integrando en los vértices superiores derechos de cada subrectángulo se obtiene una función con valores de la forma $(1 + j/N + k/2M)^3$. Cada subrectángulo tiene área $(1/N)(1/(2M))$. La suma de Riemann correspondiente es

$$\sum_{k=1}^{N}\sum_{j=1}^{M}\left(1 + \frac{j}{N} + \frac{k}{2M}\right)^3\left(\frac{1}{(2NM)}\right).$$

Para simplificar, tomamos $N = M$, con lo cual la suma toma la forma

$$S_N = \sum_{k=1}^{N}\sum_{j=1}^{N}\left(1 + \frac{j}{N} + \frac{k}{2N}\right)^3\left(\frac{1}{2N^2}\right).$$

Es conveniente usar un computador para evaluar estas sumas aun en el caso en que N no es muy grande. Se tiene la tabla

N	S_N
1	7,8125
10	3,3215
20	3,1343
50	3,0248
100	2,9888
200	2,9707

Si calculamos la integral usando integrales iteradas e integración indefinida, comprobamos que el valor correcto con dos decimales es 2,95.

En los Ejemplos 1 y 2 se hicieron las estimaciones de modo tal que el error al usar la estimación superior o inferior estaba acotado por la diferencia entre ellas. Sin embargo, en el Ejemplo 3 calculamos una sola suma de Riemann para cada subdivisión del dominio de integración. Para tener una estimación del error, a veces podemos usar el siguiente teorema.

5.2 Teorema

Sea f continua en un rectángulo cerrado y acotado R en $\mathcal{R}^n$ definido por $a_i \leq x_i \leq b_i$, $i = 1, \ldots, n$, y supongamos que f tiene derivadas parciales tales que

$$\left|\frac{\partial f}{\partial x_i}(x)\right| \leq M_i, \qquad i = 1, \ldots, n.$$

Sea G un reticulado en R cuyos rectángulos tienen ancho $h_i = (b_i - a_i)/N_i$ en la coordenada x_i, donde N_i es un entero. Entonces el error al aproximar $\int_R f\,dV$ por una suma de Riemann usando G es a lo más

$$V(R)\sum_{i=1}^{n} h_i M_i,$$

donde $V(R)$ es el volumen de R.

Demostración. Como $R = R_1 \cup \ldots \cup R_N$ y los R_k son disjuntos, tenemos que

$$\int_R f\,dV = \sum_{k=1}^{N} \int_{R_k} f\,dV.$$

Entonces,

$$\int_R f\,dV - S_N = \sum_{k=1}^{N} \int_{R_k} (f - f(x_k))\,dV,$$

ya que $f(x_k)$ es constante en R_k. Por tanto,

$$\begin{aligned}\left|\int_R f\,dV - S_N\right| &\leq \sum_{k=1}^{N} \int_{R_k} |f - f(x_k)|\,dV \\ &\leq \sum_{k=1}^{N} V(R_k) \max_{x \text{ en } R_k} |f(x) - f(x_k)|.\end{aligned}$$

Pero por el teorema del valor medio, para x e y en R_k,

$$\begin{aligned}|f(\mathbf{x}) - f(\mathbf{y})| &= |f'(\mathbf{z})(\mathbf{x} - \mathbf{y})| \\ &= \left|\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{z})(x_i - y_i)\right| \\ &\leq \sum_{i=1}^{n} h_i \left\{\max_{x \text{ en } R} \left|\frac{\partial f}{\partial x_i}(x)\right|\right\}.\end{aligned}$$

Por la desigualdad anterior,

$$\begin{aligned}\left|\int_R f\,dV - S_N\right| &\leq \sum_{k=1}^{N} V(R_k) \sum_{i=1}^{n} h_i \left\{\max_{x \text{ en } R} \left|\frac{\partial f}{\partial x_i}(x)\right|\right\} \\ &= V(R)\sum_{i=1}^{n} h_i M_i,\end{aligned}$$

como se quería demostrar.

Ejemplo 4. La integral

$$I = \int_R (1 + x + y)^3\,dx\,dy$$

del Ejemplo 3 tiene integrando $f(x, y) = (1 + x + y)^3$. Tenemos que $f_x(x, y) = f_y(x, y) = 3(1 + x + y)^2$. El rectángulo R determinado por $0 \leq x \leq 1$ y $0 \leq y \leq \frac{1}{2}$ es tal que

$$\max_{(x,y) \text{ en } R} |f_x(x, y)| = \max_{(x,y) \text{ en } R} |f_y(x, y)| = 3(\tfrac{5}{2})^2 = \tfrac{75}{4}.$$

Por tanto, $M_1 = M_2 = \frac{75}{4}$. Si subdividimos a lo largo del eje x en K partes iguales y a lo largo del eje y en L, tenemos que $h_1 = 1/k$ y $h_2 = 1/2L$. Entonces una suma de Riemann con respecto al reticulado satisfará

$$\begin{aligned}|I - S_{K,L}| &\leq (1)(\tfrac{1}{2})\left[\left(\frac{1}{K}\right)\left(\frac{75}{4}\right) + \left(\frac{1}{2L}\right)\left(\frac{75}{4}\right)\right] \\ &= \frac{75}{8}\left(\frac{1}{K} + \frac{1}{2L}\right).\end{aligned}$$

Eligiendo $L = 200$ y $K = 400$, obtenemos la siguiente cota para el error:

$$|I - S_{K,L}| < 0{,}05.$$

El caso de la integración en regiones no rectangulares es más complicado, ya que al hacer una estimación habrá un error no solo al aproximar la función que se integra, sino que también al aproximar la región de integración por un reticulado. Un dominio de integración acotado B puede ser extendido a uno rectangular simplemente extendiendo el integrando f a una función f_B que sea cero fuera de B. Sin embargo, la estimación del error dada en el Teorema 5.2 no será válida, ya que, en general, f_B será discontinua, por lo cual no tendrá las derivadas parciales requeridas. En algunos ejemplos el teorema del cambio de variable puede ser usado para transformar el dominio de integración en uno rectangular. En este caso, hay que cuidar que la transformación no complique innecesariamente al integrando.

Ejemplo 5. Sea Q la parte del disco en el primer cuadrante de $\mathcal{R}^2$ definido por $0 \leq x$, $0 \leq y$, $x^2 + y^2 \leq 1$. Usando la transformación a coordenadas polares,

$$x = u\cos v$$

$$y = u\,\mathrm{sen}\,v$$

para $0 \leq u \leq 1$ y $0 \leq v \leq \pi/2$, podemos transformar Q en el rectángulo mostrado en la Fig. 34.

Dada una integral de la forma

$$I = \int_Q f(x, y)\,dx\,dy,$$

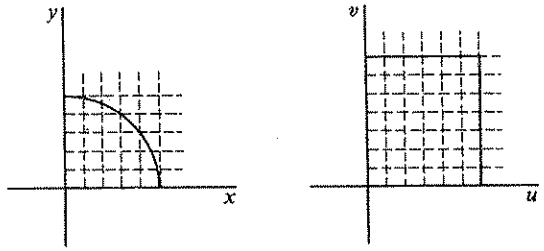


Figura 34

el teorema del cambio de variable muestra que

$$I = \int_R f(u \cos v, u \operatorname{sen} v) u \, du \, dv,$$

donde u es el determinante jacobiano de la transformación a coordenadas polares. Como R es un rectángulo, se pueden usar las estimaciones consideradas anteriormente. Por ejemplo, si $f(x, y) = (x^2 + y^2)\sqrt{x^2 + y^2}$, entonces

$$I = \int_R (u^2 \cos^2 v + u^3 \operatorname{sen}^3 v) u^2 \, du \, dv.$$

Haciendo $g(u, v) = u^4 \cos^2 v + u^5 \operatorname{sen}^3 v$, tenemos que

$$\left| \frac{\partial g}{\partial u}(u, v) \right| = |4u^3 \cos^2 v + 5u^4 \operatorname{sen}^3 v| \leq 9$$

y

$$\left| \frac{\partial g}{\partial v}(u, v) \right| = |-2u^4 \cos v \operatorname{sen} v + 3u^5 \operatorname{sen}^2 v \cos v| \leq 5$$

para $0 \leq u \leq 1$, $0 \leq v \leq \pi/2$. Ahora supondremos que $S_{N,M}$ es una suma de Riemann con respecto del reticulado formado por N subdivisiones iguales a lo largo del eje u y M a lo largo del eje v. Entonces, por el Teorema 5.1,

$$|I - S_{N,M}| \leq (1)\left(\frac{\pi}{2}\right)\left(9\left(\frac{1}{N}\right) + 5\left(\frac{\pi/2}{M}\right)\right)$$

$$\leq \frac{15}{N} + \frac{13}{M}.$$

Si tomamos $N = M = 60$, obtenemos la cota

$$|I - S_{60,60}| \leq \tfrac{15}{60} + \tfrac{13}{60} < 0{,}5.$$

para el error.

EJERCICIOS

1. Encuentre estimaciones inferiores y superiores de las siguientes integrales considerando el valor mínimo y máximo del integrando en el dominio de integración.

(a) $\displaystyle\int_0^{\pi/2} \sqrt{1 + \cos x} \, dx.$

(b) $\displaystyle\int_{-1}^{1} e^{x^2} \, dx.$

(c) $\displaystyle\int_0^1 dy \int_0^1 \cos xy \, dy.$

(d) $\displaystyle\int_{x^2+y^2 \leq 1} e^{x+y} \, dx \, dy.$

(e) $\displaystyle\int_0^1 dx \int_0^1 dy \int_0^1 \sqrt{x + y + z} \, dz.$

(f) $\displaystyle\int_0^1 dx \int_1^2 dy \int_{-1}^1 (x^2 + y^2 + z^2) \, dz.$

2. Encuentre una estimación de las integrales que siguen calculando una suma de Riemann con respecto de una subdivisión del dominio de integración en cuatro rectángulos iguales.

(a) $\displaystyle\int_0^1 (1 + x^2) \, dx.$ (b) $\displaystyle\int_0^1 dx \int_0^1 (x^2 + y^2) \, dy.$

3. Use el Teorema 5.1 para estimar el error en cada una de las aproximaciones de las integrales del Ejercicio 2.

4. Encuentre estimaciones para el error al aproximar las integrales del Ejercicio 2 si el número de subdivisiones es 10 en la parte (a) y $10^2 = 100$ en la parte (b). Use el Teorema 5.1.

5. Use un computador para hacer estimaciones de las sumas de Riemann del Ejercicio 4.

6. Si una integral se aproxima por una suma de Riemann con los valores de la función tomada en el centro de cada rectángulo, entonces se dice que la integral ha sido aproximada por la **regla del punto medio.** Se puede demostrar que en este caso la estimación del error dada por el Teorema 5.1 es

$$\frac{V(R)}{24} \sum_{i=1}^{n} h_i^2 M_i^{(2)},$$

donde $M_i^{(2)} \geq |(\partial^2 f/\partial x_i^2)(x)|$ para $i = 1, 2, \ldots, n$. Si las integrales del Ejercicio 2 se estiman usando la regla del punto medio con 10 subdivisiones iguales en la parte (a) y $10^2 = 100$ en la parte (b), encuentre una cota para el error.

7. Transforme la integral

$$\int_{x^2+y^2 \geq 1} (x + y^2) \, dx \, dy.$$

en una sobre un rectángulo. Entonces use el Teorema 5.1 para estimar el error al aproximarla por una suma de Riemann con $10^2 = 100$ subdivisiones iguales.

6. Integración numérica

Las estimaciones de la sección anterior son bastante menos precisas de lo que se requiere generalmente en los cálculos numéricos. En esta sección queremos mostrar cómo obtener una precisión mayor. Teóricamente, el Teorema 5.2 da una precisión tan grande como se desee. Sin embargo, para alcanzarla se puede necesitar un tiempo demasiado largo para hacer los cálculos. Además, como hay que hacer una gran cantidad de operaciones aritméticas, la acumulación de los

errores de redondeo puede ser demasiado grande. Para eliminar estos problemas necesitamos métodos que den una buena aproximación sin muchos cálculos aritméticos. Uno de estos métodos es la regla de Simpson, que describimos primero para integrales unidimensionales para aplicarla después a integrales múltiples.

La motivación para la regla de Simpson está en la observación que, para un polinomio cuadrático $q(x) = Ax^2 + Bx + C$,

$$\int_a^b q(x)\,dx = \frac{b-a}{6}\left[q(a) + 4q\left(\frac{a+b}{2}\right) + q(b)\right]. \tag{1}$$

En el Ejemplo 5 de la Sec. 2 del Cap. 2 se indicó cómo demostrar esta fórmula. Para usar la Ec. (1) para aproximar $\int_a^b f(x)\,dx$, hacemos

$$x_0 = a, \qquad x_1 = \frac{(a+b)}{2}, \qquad x_2 = b$$

y encontramos un polinomio cuadrático q tal que

$$q(x_0) = f(x_0), \qquad q(x_1) = f(x_1), \qquad q(x_2) = f(x_2). \tag{2}$$

[Como $q(x)$ tiene la forma $Ax^2 + Bx + C$, para ciertos A, B, C, para obtener q basta resolver en A, B, C el sistema de ecuaciones $Ax_k^2 + Bx_k + C = f(x_2)$, $k = 1, 2, 3$. Sin embargo, en la práctica, no es necesario hacer esto.] Entonces, tomamos la aproximación

$$\int_a^b f(x)\,dx \approx \int_a^b q(x)\,dx.$$

Como por la Ec. (2), f y q son iguales en x_0, x_1 y x_2, usando la Ec. (1) obtenemos que

6.1 $$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f(x_0) + 4f(x_1) + f(x_2)\right].$$

Esta última forma se llama la **aproximación de Simpson por tres puntos**. La

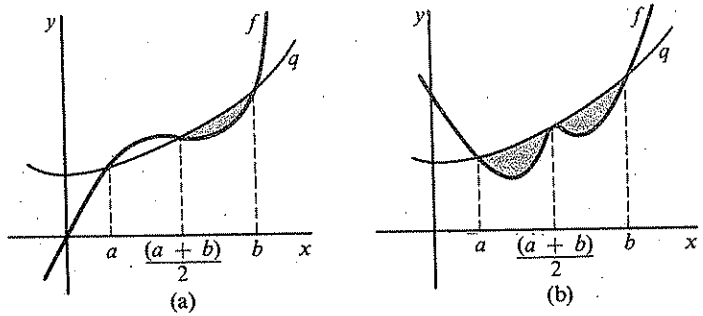
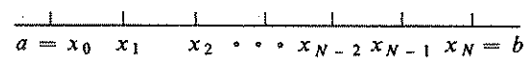


Figura 35

Fig. 35 muestra que la aproximación puede ser buena o mala dependiendo de f. En la Fig. 35(a), el gráfico de q está cerca del de f sobre $[a, b]$; además, hay cancelación entre el área sombreada y la no sombreada, ya que q está parcialmente sobre y debajo de f. En la Fig. 35(b), los gráficos no están muy cerca y no hay cancelación.

Para mejorar la aproximación, subdividimos $[a, b]$ en un número *par* de intervalos iguales tomando puntos $x_0, x_1, \ldots, x_N$, N par.



Como

$$\int_a^b f(x)\,dx = \int_{x_0}^{x_2} f(x)\,dx + \int_{x_2}^{x_4} f(x)\,dx + \ldots + \int_{x_{N-2}}^{x_N} f(x)\,dx,$$

podemos aplicar la regla de los tres puntos $N/2$ veces obteniendo

$$\begin{aligned}\int_a^b f(x)\,dx \approx{}& \frac{b-a}{6(N/2)}\left[f(x_0) + 4f(x_1) + f(x_2)\right]\\ &+ \frac{b-a}{6(N/2)}\left[f(x_2) + 4f(x_3) + f(x_4)\right]\\ &\cdot\\ &\cdot\\ &\cdot\\ &+ \frac{b-a}{6(N/2)}\left[f(x_{N-2}) + 4f(x_{N-1}) + f(x_N)\right].\end{aligned}$$

ombinando términos, tenemos que

6.2 $$\int_a^b f(x)\,dx \approx \frac{b-a}{3N}\left[f(x_0) + 4f(x_1) + 2f(x_2) + \ldots + 4f(x_{N-1}) + f(x_N)\right].$$

La Ec. 6.2 se llama la **regla de Simpson**.

Ejemplo 1. Aplicando la regla de Simpson a $I = \int_0^1 e^{x^2}\,dx$ para valores pequeños de N, da las siguientes aproximaciones S_N:

$$\frac{1-0}{3N}\left[e^0 + 4e^{1/N} + 2e^{2/N} + \ldots + 4e^{(N-1)/N} + e^1\right].$$

N	S_N
2	1,4757
4	1,4637
10	1,4627
20	1,4626

La Ec. 6.2 no dice nada acerca de la precisión dada por la regla de Simpson. Esta información está contenida en el siguiente teorema.

6.3 Teorema

Sea f una función continua en $[a, b]$ cuatro veces diferenciable en (a, b). Sea M un número tal que $|f^{(4)}(x)| \leq M$ para todo x en (a, b). Entonces el error en la aproximación de Simpson

$$\int_a^b f(x)\,dx \approx \frac{b-a}{3N}[f(x_0) + 4f(x_1) + 2f(x_2) + \ldots + 4f(x_{N-1}) + f(x_N)],$$

usando N intervalos (N par), es a lo más

$$\frac{(b-a)^5 M}{180N^4}.$$

Demostración. Primero reducimos la demostración al caso $N = 2$. Para ello dividimos $[a, b]$ en $N/2$ *pares* de intervalos de largo $(b-a)/(N/2)$. Si podemos demostrar que el error en cada par es a lo más

$$\frac{((b-a)/N)^5 M}{90},$$

entonces el error total es a lo más $N/2$ veces eso, es decir,

$$\frac{(b-a)^5 M}{180N^4}.$$

Por tanto, para demostrar el teorema solo necesitamos considerar la regla de los tres puntos en $[a, b]$ y demostrar que

$$\left| \int_a^b f(x)\,dx - \frac{h}{3}[f(a) + 4f(c) + f(b)] \right| \leq \frac{h^5 M}{90},$$

donde $h = (b-a)/2$ y $c = (b+a)/2$. Primero tomemos

$$E(t) = \int_{c-t}^{c+t} f(x)\,dx - \frac{t}{3}[f(c+t) + 4f(c) + f(c-t)]$$

y

$$F(t) = E(t) - \frac{t^5}{h^5}E(h).$$

Un cálculo sencillo muestra que

$$F'(t) = \tfrac{2}{3}[f(c+t) - 2f(c) + f(c-t)] - \tfrac{1}{3}t[f'(c+t) - f'(c-t)] - \frac{5t^4}{h^5}E(h),$$

$$F''(t) = \tfrac{1}{3}[f'(c+t) - f'(c-t)] - \tfrac{1}{3}t[f''(c+t) + f''(c-t)] - \frac{20t^3}{h^4}E(h),$$

$$F'''(t) = -\tfrac{1}{3}t[f'''(c+t) - f'''(c-t)] - \frac{60t^2}{h^5}E(h).$$

Ahora aplicamos el teorema del valor medio a F en $t = 0$ y $t = h$. Pero como $F(0) = F(h) = 0$, concluimos que $F'(t_1) = 0$ para algún t_1 entre 0 y h. Como $F'(0) = 0$, podemos aplicar el teorema del valor medio a F' en $t = 0$ y $t = t_1$, concluyendo que $F''(t_2) = 0$ para algún t_2 entre 0 y t_1. Como $F''(0) = 0$, concluimos del mismo modo que $F'''(t_3) = 0$ para algún t_3 entre 0 y t_2. Esto implica que

$$E(h) = \frac{h^5}{180t_3}[f'''(c+t_3) - f'''(c-t_3)]. \tag{3}$$

Usando nuevamente el teorema del valor medio, tenemos que

$$f'''(c+t_3) - f'''(c-t_3) = 2t_3 f^{(4)}(t_4) \tag{4}$$

para algún t_4 entre $c - t_3$ y $c + t_3$. Como $0 < t_3 < h$, el número t_4 está en (a, b). Usando (3) y (4), tenemos que

$$E(h) = \frac{h^5}{90}f^{(4)}(t_4).$$

lo cual termina la demostración.

Ejemplo 2. Aplicamos el teorema anterior a

$$\int_0^1 e^{x^2}\,dx$$

Tenemos que $f^{(4)}(x) = 12e^{x^2} + 48x^2e^{x^2} + 16x^4e^{x^2}$, de modo que en $[0, 1]$ se tiene que $|f^{(4)}(x)| \leq 12e + 48e + 16e < 228$. Por tanto, el error al remplazar la integral por la aproximación de Simpson usando N puntos es a lo más

$$\frac{(1-0)^5 228}{180N^4} < \frac{1{,}3}{N^4}.$$

Para $N = 10$, el error es a lo más 0,00013 y para $N = 100$, a lo más 0,000000013.

Para ver cómo se aplica la regla de Simpson a una integral múltiple sobre un rectángulo, aplicamos varias veces la Ec. 6.2 a la integral iterada en algún orden. Por ejemplo, si

$$\int_R f(x, y)\,dx\,dy = \int_c^d dy \int_a^b f(x, y)\,dx,$$

consideramos para cada y en $[c, d]$ la aproximación

$$F(y) = \int_a^b f(x, y)\,dy \approx \frac{b-a}{3N}\sum_{j=0}^{J} A_j f(x_j, y), \tag{5}$$

donde los números A_j se obtienen como en la regla de Simpson: $A_0 = 1$, $A_1 = 4$, $A_2 = 2, \ldots, A_{J-1} = 4$, $A_J = 1$ y J es par. Los puntos x_j están uniformemente espaciados en $a + j(b - a)/J$, para $j = 0, 1, \ldots, J$. Aplicamos la regla de Simpson a la integral de F con K intervalos en $[c, d]$, obteniendo

$$\int_c^d F(y)\,dy \approx \frac{d-c}{3K}\sum_{k=0}^{K} B_k F(y_k), \tag{6}$$

donde $B_0 = 1$, $B_1 = 4, \ldots, B_{K-1} = 4$, $B_K = 1$, como de costumbre. Bajo riesgo de aumentar el error, sustituimos la aproximación de la Ec. (5) en la de (6), obteniendo

$$\int_c^d dy \int_a^b f(x, y)\,dx \approx \frac{(b-a)}{3J}\,\frac{(d-c)}{3K}\sum_{k=0}^{K} B_k \sum_{j=0}^{J} A_j f(x_j, y_k)$$

es decir,

$$\textbf{6.4} \qquad \int_R f(x, y)\,dx\,dy \approx \frac{A(R)}{9JK}\sum_{k=0}^{K}\sum_{j=0}^{J} A_j B_k f(x_j, y_k),$$

donde $x_j = a + j(b - a)/J$ y $j_k = c + k(d - c)/K$. Los productos $A_j B_k$ son simplemente los productos de los coeficientes habituales de Simpson. Por ejemplo, si $A_3 = 4$ y $B_2 = 2$, $A_3 B_2 = 8$.

Ejemplo 3. Consideremos

$$\int_R (x - y)^4\,dx\,dy = \int_1^2 dy \int_1^3 (x - y)^4\,dx.$$

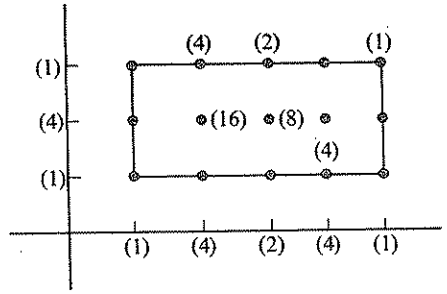


Figura 36

Tomando $J = 4$ y $K = 2$, encontramos una aproximación evaluando el integrando en los puntos indicados en la Fig. 36. Los coeficientes asociados, $A_j B_k$, se muestran en varios casos.

De la Ec. 6.4, se tiene que

$$\int_R (x - y)^4\,dx\,dy \approx \frac{2}{9(2)(4)}\sum_{k=0}^{2}\sum_{j=0}^{4} A_j B_k (x_j - y_k)^4,$$

donde $x_j = 1 + j/2$ e $y_k = 1 + k/2$. La suma de la derecha tiene 15 sumandos. Usando un computador, obtenemos $S_{4,2} = 2{,}13542$ para la aproximación de Simpson con $J = 4$ y $K = 2$. Otros valores son los siguientes:

J, K	$S_{J,K}$
8, 4	2,13542
16, 8	2,13346
32, 16	2,13334
64, 32	2,13333

Sucede que la integral que hemos considerado puede ser evaluada por integración indefinida iterada y su valor es $\frac{32}{15}$, que coincide con $S_{32,16}$ con el número de decimales considerado.

Para aproximar una integral triple de la forma

$$\int_R f(x, y, z)\,dV = \int_{a_3}^{b_3} dz \int_{a_2}^{b_2} dy \int_{a_1}^{b_1} f(x, y, z)\,dx,$$

aplicamos tres veces la regla de Simpson en una variable obteniendo así la regla de Simpson tridimensional:

$$\textbf{6.5} \qquad \int_R f(x, y, z)\,dV \approx \frac{V(R)}{27JKL}\sum_{l=0}^{L}\sum_{k=0}^{K}\sum_{j=0}^{J} A_j B_k C_l f(x_j, y_k, z_l),$$

donde $x_j = a_1 + j(b_1 - a_1)/J$, $y_k = a_2 + k(b_2 - a_2)/K$ y $z_l = a_3 + l(b_3 - a_3)/L$. Nuevamente, los A, B y C son los coeficientes usuales de Simpson, de modo que, por ejemplo, si $A_2 = 2$, $B_3 = 4$ y $C_3 = 4$, tenemos que $A_2 B_3 C_3 = 32$. Nótese que, en el caso $J = K = L = 2$, la suma en 6.5 tiene 27 términos, de modo que es deseable tener un computador para hacer los cálculos aritméticos.

Ejemplo 4. Para aproximar la integral

$$\int_{1,1}^{2,3} dz \int_{1,5}^{1,8} dy \int_{1}^{3,6} \log(xyz)\,dx,$$

aplicamos la Ec. 6.4 al rectángulo sólido $1 \le x \le 3{,}6$, $1{,}5 \le y \le 1{,}8$, $1{,}1 \le z \le 2{,}3$. Las aristas de este rectángulo tienen largos 2,6, 0,3 y 1,2, respectivamente, de modo que elegimos J, K y L en esa proporción. Como primera aproximación usamos J, K y L en $L = 8$. El número de términos en la suma $S_{16,2,8}$ es $(17)(3)(9) = 459$.

Tenemos que

$$S_{16,2,8} = \frac{(3{,}6 - 1)(1{,}8 - 1{,}5)(2{,}3 - 1{,}1)}{27(16)(2)(8)}\sum_{l=0}^{8}\sum_{k=0}^{2}\sum_{j=0}^{16} A_j B_k C_l \log(x_j y_k z_l),$$

donde $x_j = 1 + j(3{,}6 - 1)/8$, $y_k = 1{,}5 + k(1{,}8 - 1{,}5)/2$ y $z_l = 1{,}1 + l(2{,}3 - 1{,}1)/8$. Efectuando los cálculos, se tiene que $S_{16,2,8}$ es aproximadamente 1,66797.

Para mejorar la precisión en el Ejemplo 4, podemos refinar la subdivisión. Sin embargo, como los cálculos son demasiado complicados, es conveniente tener una estimación del error como la de 6.3 para usar como guía al elegir el reticulado.

6.6 Teorema

Sea f continua en un rectángulo cerrado y acotado R en $\mathcal{R}^n$, con derivadas parciales de cuarto orden que satisfacen

$$\left|\frac{\partial^4 f}{\partial x_i^4}(x)\right| \leq M_i, \qquad i = 1, \ldots, n.$$

Supongamos que dividimos R mediante un reticulado G en rectángulos de ancho

$$h_i = \frac{b_i - a_i}{N_i}$$

en la coordenada x_i, donde cada N_i es un entero par. Entonces el error al tomar la aproximación de Simpson, basado en G, es a lo sumo

$$\frac{V(R)}{180}\sum_{i=1}^{n} h_i^4 M_i.$$

Demostración. El teorema tiene a 6.3 como caso especial en que $n = 1$ y se demuestra reduciéndolo a una aplicación repetida de 6.3. El caso $n = 2$ contiene la idea fundamental de la demostración, puesto que es el paso esencial en una inducción. Haciendo

$$E = \int_c^d dy \int_a^b f(x, z)\, dy - \frac{(b-a)(d-c)}{9JK}\sum_{j=0}^{J} A_j \sum_{k=0}^{K} B_k f(x_j, y_k),$$

sumamos y restamos un término intermedio para obtener

$$\begin{aligned} E = {} & \int_c^d dy \int_a^b f(x, y)\, dy - \frac{b-a}{3J}\sum_{j=0}^{J} A_j \int_c^d f(x_j, y)\, dy \\ & + \frac{b-a}{3J}\sum_{j=0}^{J} A_j \int_c^d f(x_j, y)\, dy \\ & - \frac{(b-a)(d-c)}{9JK}\sum_{j=0}^{J} A_j \sum_{k=0}^{K} B_k f(x_j, y_k). \end{aligned}$$

Usando la desigualdad triangular e intercambiando la suma y la integración, se tiene que

$$\begin{aligned} |E| \leq {} & \left| \int_c^d \left[\int_a^b f(x, y)\, dx - \frac{b-a}{3J}\sum_{j=0}^{J} A_j f(x_j, y) \right] dy \right| \\ & + \left| \frac{b-a}{3J}\sum_{j=0}^{J} A_j \left[\int_c^d f(x_j, y)\, dy - \frac{d-c}{3K}\sum_{k=0}^{K} B_k f(x_j, y_k) \right] \right|. \end{aligned}$$

Ahora aplicamos el caso de una variable 6.3 a cada una de las dos expresiones en paréntesis cuadrados. Entonces

$$|E| \leq \int_c^d \left[\frac{(b-a)^5}{180J^4} M_1 \right] dy + \frac{b-a}{3J}\sum_{j=0}^{J} A_j \left[\frac{(d-c)^5}{180K^4} M_2 \right].$$

Pero ahora las expresiones en paréntesis cuadrados son independientes de y y j, respectivamente. Además, es fácil verificar que $\sum_{j=0}^{J} A_j = 3J$. Por tanto,

$$\begin{aligned} |E| &\leq (d-c)\frac{(b-a)^5}{180J^4} M_1 + (b-a)\frac{(d-c)^5}{180K^4} M_2 \\ &= \frac{(d-c)(b-a)}{180}\left[\frac{(b-a)^4}{J^4} M_1 + \frac{(d-c)^4}{K^4} M_2 \right], \end{aligned}$$

que es lo que queríamos demostrar.

Ejemplo 5. Volviendo a la integral

$$\int_{1,1}^{2,3} dz \int_{1,5}^{1,8} dy \int_{1}^{3,6} \log(xyz)\, dx$$

del Ejemplo 4, encontramos que

$$V(R) = (2{,}3 - 1{,}1)(1{,}8 - 1{,}5)(3{,}6 - 1) < 1.$$

Además, como el integrando es $f(x, y, z) = \log(xyz)$, tenemos que

$$\frac{\partial^4 f}{\partial x^4}(x, y, z) = -6x^{-4}$$

$$\frac{\partial^4 f}{\partial y^4}(x, y, z) = -6y^{-4}$$

$$\frac{\partial^4 f}{\partial z^4}(x, y, z) = -6z^{-4}.$$

Como x, y y z son todos mayores o iguales que 1 en R, las derivadas parciales de cuarto orden tienen valor absoluto a lo más 6. Por tanto, podemos elegir $M_1 = M_2 = M_3 = 6$. Los números h_1, h_2 y h_3 serán estimados más cuidadosamente, ya que deben ser elevados a la cuarta potencia. Tomamos

$$h_1 = \frac{3{,}6 - 1}{J} = \frac{2{,}6}{J}$$

$$h_2 = \frac{1{,}8 - 1{,}5}{L} = \frac{0{,}3}{K}$$

$$h_3 = \frac{2{,}3 - 1{,}1}{L} = \frac{1{,}2}{L}.$$

Entonces el Teorema 6.6 nos da una estimación del error

$$E < \frac{1}{180}\left(\left(\frac{2,6}{J}\right)^4(6) + \left(\frac{0,3}{K}\right)^4(6) + \left(\frac{1,2}{L}\right)^4(6)\right).$$

Si calculamos este número para $J = 16$, $K = 2$ y $L = 8$, tenemos que $E < 0,000029$. Recordemos que en el Ejemplo 4 calculamos el valor aproximado de la integral usando la regla de Simpson y obtuvimos un valor 1,66797. La estimación del error nos da una latitud posible de $\pm 0,000029$. En otras palabras, el Teorema 6.6 indica que el valor verdadero de la integral está entre 1,66794 y 1,66800.

EJERCICIOS

1. Al usar la regla de Simpson para aproximar $\int_0^6 x \operatorname{sen} x\, dx$, ¿cuál es el número razonable más pequeño de subintervalos que se deben usar para emplear ventajosamente la forma del gráfico de $x \operatorname{sen} x$?
2. Use la regla de los tres puntos para estimar cada una de las integrales siguientes:
   (a) $\int_1^3 x^4\, dx$. (b) $\int_0^2 \operatorname{sen} x\, dx$. (c) $\int_0^2 e^{-x^2}\, dx$.
3. Use la regla de Simpson con cuatro intervalos para estimar cada una de las siguientes integrales:
   (a) $\int_0^1 x \operatorname{sen} x\, dx$, (b) $\int_1^{1,4} \log \operatorname{sen} x\, dx$. (c) $\int_{-1}^0 e^x\, dx$.
4. Usando un computador aplique la regla de Simpson con 100 intervalos para estimar las siguientes integrales:
   (a) $\int_0^2 e^{x^2}\, dx$. (b) $\int_0^{2,1} \frac{dx}{x^3 + 1}$. (c) $\int_{1,1}^{2,3} \frac{dx}{x}$.
5. Use el Teorema 6.3 para estimar los errores en las estimaciones del Ejercicio 4.
6. Aplique la Fórmula 6.4 con $J = K = 2$ para estimar las integrales
   (a) $\int_0^1 \int_1^2 (x + y)\, dx\, dy$. (b) $\int_{-1}^1 dx \int_0^1 (x - y)\, dy$.
7. Use un computador y la Fórmula 6.4 con $J = K = 8$ para estimar las integrales
   (a) $\int_0^1 \int_1^2 (x + e^y)^5\, dx\, dy$. (b) $\int_{1,1}^{2,1} dx \int_0^{1,1} \log (x + y)\, dy$.
8. Use el Teorema 6.6 para estimar los errores en las aproximaciones del Ejercicio 6.
9. Use el Teorema 6.6 para estimar los errores en las aproximaciones del Ejercicio 7.
10. ¿Qué valores de J y K deben ser tomados en la Fórmula 6.4 para que la aproximación de Simpson de
$$\int_0^1 dx \int_0^2 e^{xy}\, dy$$

tenga un error no mayor de 0,0001? Recuerde que J y K deben ser enteros pares.

11. ¿Qué valores de J, K y L hay que elegir en la Fórmula 6.5 para que la aproximación de Simpson de
$$\int_0^1 dx \int_0^{1,1} dy \int_0^{1,2} e^{xyz}\, dz$$
tenga un error no mayor de 0,0001?
12. Sea A la región anular en $\mathfrak{R}^2$ definida por $1 \leq x^2 + y^2 \leq 4$; es decir, la región comprendida entre los círculos de centro (0, 0) y radio 1 y 2. La integral
$$\int_A (x^2 + 2y^2)\, dx\, dy$$
no puede ser bien aproximada directamente por la regla de Simpson ya que A no es rectangular.
   (a) Transforme la integral mediante el teorema del cambio de variables en una integral sobre un rectángulo.
   (b) Aproxime la integral encontrada en la parte (a) con un error no mayor de 0,005.

7

Teoría de los campos vectoriales

1. El teorema de Green

El teorema fundamental del cálculo dice que si f' es integrable para $a \le t \le b$, entonces

$$\int_a^b f'(t)\,dt = f(b) - f(a). \tag{1}$$

En la Sec. 2 del Cap. 4 este teorema ha sido extendido a las integrales de línea de un gradiente ∇f mediante la ecuación

$$\int_a^b \nabla f(\mathbf{x}) \cdot d\mathbf{x} = f(\mathbf{b}) - f(\mathbf{a}). \tag{2}$$

Los teoremas principales de este capítulo son también variaciones de la idea de que la integral de algún tipo de derivada de una función puede ser evaluada usando los valores de la función en un conjunto de dimensión más baja. Empezamos con la versión conocida como teorema de Green.

Sea D una región del plano cuya frontera es una curva γ parametrizada por una función g de tal manera que, cuando t crece de a a b, $g(t)$ recorre γ una vez en la dirección contraria a la de los punteros del reloj. La Fig. 1 muestra un ejemplo. Si F_1 y F_2 son funciones reales definidas en D y su frontera, entonces el teorema de Green dice que

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1\,dx_2 = \int_\gamma F_1\,dx_1 + F_2\,dx_2, \tag{3}$$

bajo condiciones adecuadas de diferenciabilidad. El requisito de que γ sea recorrida en el sentido contrario al de los punteros del reloj es el análogo del hecho

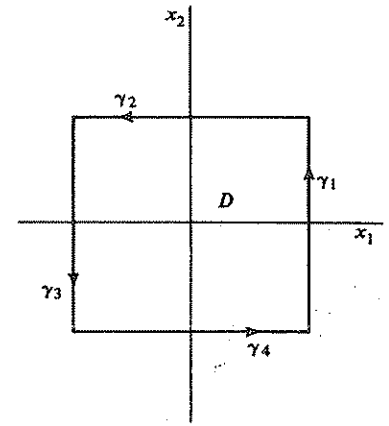


Figura 1

de que en las Ecs. (1) y (2) las diferencias que aparecen a la derecha deben ser tomadas en el orden apropiado. La analogía de la Ec. (3) con las Ecs. (1) y (2) puede ser fortalecida si consideramos el integrando $(\partial F_2/\partial x_1 - \partial F_1/\partial x_2)$ como una especie de derivada del campo vectorial $F = (F_1, F_2)$. En la Sec. 8 se dará una justificación de este punto de vista.

Ejemplo 1. Supongamos que D es el cuadrado definido por $-1 \le x_1 \le 1$, $-1 \le x_2 \le 1$ y sean F_1 y F_2 definidas en D por $F_1(x_1, x_2) = -x_2 e^{x_1}$ y $F_2(x_1, x_2) = x_1 e^{x_2}$. Entonces

$$\frac{\partial F_2}{\partial x_1}(x_1, x_2) - \frac{\partial F_1}{\partial x_2}(x_1, x_2) = e^{x_2} + e^{x_1};$$

de modo que

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1\,dx_2 = \int_{-1}^{1} dx_1 \int_{-1}^{1} (e^{x_2} + e^{x_1})\,dx_2$$
$$= 4\left(e - \frac{1}{e}\right).$$

La curva frontera γ puede ser parametrizada en cuatro partes γ_i, $i = 1, 2, 3, 4$, por

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 \\ t \end{pmatrix} \\ \begin{pmatrix} -t \\ 1 \end{pmatrix} \\ \begin{pmatrix} -1 \\ -t \end{pmatrix} \\ \begin{pmatrix} t \\ -1 \end{pmatrix} \end{cases}, \qquad -1 \le t \le 1.$$

Nótese que γ es recorrido en el sentido opuesto al de los punteros del reloj, tal como lo muestra la Fig. 1. En el primer lado del cuadrado tenemos que

$$\begin{aligned}\int_{\gamma_1} F_1\,dx_1 + F_2\,dx_2 &= \int_{\gamma} -x_2e^{x_1}\,dx_1 + x_1e^{x_2}\,dx_2\\ &= \int_{-1}^{1}\left[(-te)\frac{dx_1}{dt} + e^t\frac{dx_2}{dt}\right]dt\\ &= \int_{-1}^{1} e^t\,dt = e - \frac{1}{e}.\end{aligned}$$

Análogamente, las integrales sobre los otros lados del cuadrado, también son iguales a $(e - 1/e)$, de modo que

$$\int_{\gamma} F_1\,dx_1 + F_2\,dx_2 = 4\left(e - \frac{1}{e}\right).$$

Por tanto, la Ec. (3) ha sido verificada en este caso particular.

Al calcular una integral de línea, toda parametrización puede ser remplazada por una equivalente, obteniéndose el mismo valor para la integral. En el ejemplo anterior, la curva frontera γ ha sido parametrizada de una manera muy sencilla. Si la frontera está dada solo como un conjunto, es necesario elegir una parametrización y veremos que, si queremos aplicar el teorema de Green, la elección de parametrización debe ser tal que la curva sea recorrida una sola vez y en la dirección adecuada.

La importancia de la dirección del recorrido está en que si la cambiamos, el signo de la integral de línea también cambia. Así, si γ está parametrizada por $g(t)$ para $a \le t \le b$, podemos denotar por γ^- la curva parametrizada por $g^-(t) = g(a + b - t)$ para $a \le t \le b$. Es claro que γ^- es, como conjunto, igual a γ, pero es recorrida en la dirección opuesta, es decir, de $g(b)$ a $g(a)$. Entonces, como $g^{-1}(t) = -g'(a + b - t)$, tenemos que

$$\int_{\gamma^-} F\cdot d\mathbf{x} = -\int_a^b F(g(a + b - t))\cdot g'(a + b - t)\,dt.$$

Haciendo el cambio de variable $t = a + b - u$, se tiene que

$$\begin{aligned}\int_{\gamma^-} F\cdot d\mathbf{x} &= -\int_a^b F(g(u))\cdot g'(u)\,du\\ &= -\int_{\gamma} F\cdot d\mathbf{x}\end{aligned}$$

Esto demuestra la importante fórmula

$$\int_{\gamma^-} F\cdot d\mathbf{x} = -\int_{\gamma} F\cdot d\mathbf{x}. \tag{1.1}$$

El teorema de Green puede ser demostrado muy fácilmente para regiones D tales que su frontera γ es cruzada a lo más dos veces por cualquier recta paralela a un eje de coordenadas. Las regiones de este tipo se llaman **simples**. Por tanto, una recta coordenada interseca la frontera de una región simple en un segmento de recta o en a lo más dos puntos. Nótese que si se demuestra el teorema para regiones simples usando la Ec. 1.1, lo podemos demostrar para las uniones finitas de regiones simples. La Fig. 2 muestra varias regiones que son uniones finitas de regiones simples, de las cuales solo D_1 es simple. La frontera D_2 se muestra de tal manera, que no siempre tiene la dirección contraria a la de los punteros del reloj. Se ha elegido la dirección de recorrido, de modo que, al recorrer la curva, el interior de la región está siempre a la izquierda. Para regiones acotadas con una sola curva frontera, las dos descripciones de la orientación de una frontera coinciden.

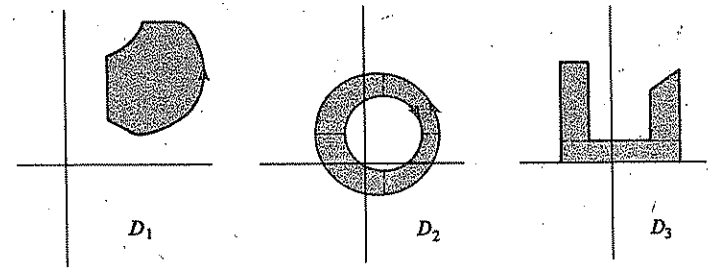


Figura 2

1.2 Teorema de Green

Sea D una región plana acotada, unión finita de regiones simples, cada una de las cuales tiene una frontera formada por una curva lisa por pedazos. Sean F_1 y F_2 dos funciones reales continuamente diferenciables definidas en D y en γ (la frontera de D). Entonces,

$$\int_D\left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right)dx_1\,dx_2 = \int_{\gamma} F_1\,dx_1 + F_2\,dx_2,$$

donde γ está parametrizada, de modo que ella quede recorrida una sola vez con D a la izquierda.

Demostración. Consideremos primero el caso en que D es una región simple con frontera γ parametrizada por

$$\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}g_1(t)\\ g_2(t)\end{pmatrix}, \qquad a \le t \le b.$$

Como

$$\int_{\gamma} F_1\,dx_1 + F_2\,dx_2 = \int_{\gamma} F_1\,dx_1 + \int_{\gamma} F_2\,dx_2,$$

podemos trabajar separadamente con cada término de la derecha de la ecuación. Tenemos que

$$\int_\gamma F_1(x_1, x_2)\, dx_1 = \int_a^b F_1(g_1(t), g_2(t))g_1'(t)\, dt.$$

Tal como lo muestra la Fig. 3, la curva γ está formada por los gráficos de dos funciones $u(x_1)$ y $v(x_1)$ y, posiblemente, por uno o dos segmentos de recta verticales. En un segmento vertical, g_1 es constante, de modo que en esa parte $g_1' = 0$. Aplicamos el cambio de variables $x_1 = g_1(t)$ a las partes restantes de γ, de modo que en la curva superior, $g_2(t) = u(x_1)$ y en la inferior, $g_2(t) = v(x_1)$. Por tanto,

$$\int_\gamma F_1(x_1, x_2)\, dx_1 = \int_\beta^\alpha F_1(x_1, u(x_1))\, dx_1 + \int_\alpha^\beta F_1(x_1, v(x_1))\, dx_1,$$

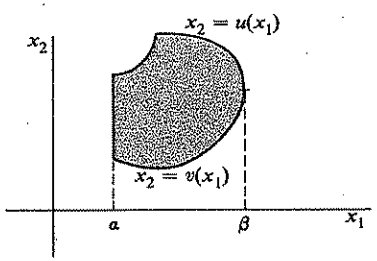


Figura 3

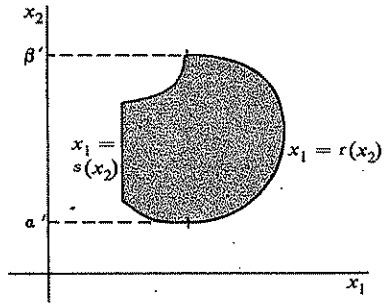


Figura 4

donde la integral de β a α aparece debido a que el gráfico de u está recorrido de derecha a izquierda. Invirtiendo los límites en la primera integral tenemo que

$$\begin{aligned}\int_\gamma F_1(x_1, x_2)\, dx &= \int_\alpha^\beta [-F_1(x_1, u(x_1)) + F_1(x_1, v(x_1))]\, dx_1 \\ &= \int_a^\beta \left[-\int_{v(x_1)}^{u(x_1)} \frac{\partial F_1}{\partial x_2}(x_1, x_2)\, dx_2 \right] dx_1 \\ &= \int_D -\frac{\partial F_1}{\partial x_2}\, dx_1\, dx_2.\end{aligned}$$

Una demostración análoga usando la Fig. 4, demuestra que

$$\int_\gamma F_2(x_1, x_2)\, dx_2 = \int_D \frac{\partial F_2}{\partial x_1}\, dx_1\, dx_2.$$

Combinando esta ecuación con las anteriores, tenemos el teorema de Green para las regiones simples.

Extendemos el resultado a una unión finita $D = D_1 \cup \ldots \cup D_K$ de regiones simples, cada una de las cuales tiene una frontera, lisa por pedazos γ_k, $k = 1, \ldots, K$. Aplicando el teorema de Green a cada región simple D_k, se tiene

$$\int_{D_k} \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1\, dx_2 = \int_{\gamma_k} F_1\, dx_1 + F_2\, dx_2$$

La suma de las integrales es una integral sobre D, de modo que

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1\, dx_2 = \int_{\gamma_1} F_1\, dx_1 + F_2\, dx_2 + \ldots + \int_{\gamma_K} F_1\, dx_1 + F_2\, dx_2.$$

Ahora bien, la frontera de D está formada por pedazos tomados de varias de las curvas γ_k. Además, puede haber pedazos de las curvas γ_k que no son parte de γ y que son parte de la frontera común de dos de las regiones simples. La Fig. 5 ilustra esta situación.

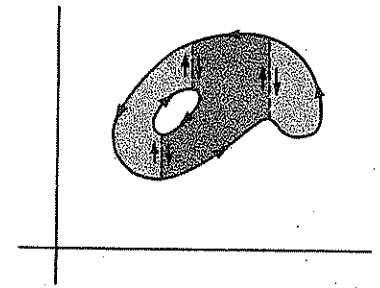


Figura 5

Un pedazo común, δ, de frontera será recorrido en direcciones opuestas cuando se la considere como parte de la frontera de cada una de las regiones. Pero por la Ec. 1.1, tenemos que

$$\int_\delta F_1\, dx_1 + F_2\, dx_2 + \int_{-\delta} F_1\, dx_1 + F_2\, dx_2 = 0.$$

Entonces, solo los pedazos de las curvas γ_k que forman parte de γ contribuyen a la integral $\int_\gamma F_1\, dx_1 + F_2\, dx_2$. Las otras partes se cancelan y se tiene que

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1\, dx_2 = \int_\gamma F_1\, dx_1 + F_2\, dx_2.$$

Esto completa la demostración del teorema de Green.

La última parte de la demostración recién dada extiende el teorema de Green de regiones simples a regiones del tipo indicado en la Fig. 6. Esta extensión tiene una consecuencia importante para las integrales de línea $\int F_1\, dx_1 + F_2\, dx_2$

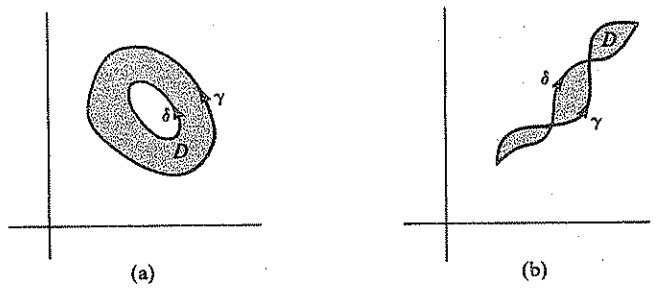


Figura 6

sobre dos curvas γ y δ, cuando F_1 y F_2 son continuamente diferenciables en la región entre γ y δ y sobre las curvas frontera. En la Fig. 6(a), las curvas están trazadas en la misma dirección (contraria a la de los punteros del reloj) y en la Fig. 6(b) las curvas van en la misma dirección de un punto a otro. Si sucede que la ecuación

$$\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} = 0 \tag{4}$$

es válida en D, entonces podemos deducir que

$$\int_\gamma F_1\,dx_1 + F_2\,dx_2 = \int_\delta F_1\,dx_1 + F_2\,dx_2.$$

Los dos ejemplos siguientes ilustran este principio.

Ejemplo 2. Sean F_1 y F_2 definidas por

$$F_1(x_1, x_2) = \frac{-x_2}{x_1^2 + x_2^2}, \qquad F_2(x_1, x_2) = \frac{x_1}{x_1^2 + x_2^2},$$

para $(x_1, x_2) \neq (0, 0)$. Un cálculo directo muestra que estas funciones satisfacen la Ec. (4). Si γ es la elipse definida por

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2\cos t \\ 3\,\mathrm{sen}\,t \end{pmatrix}, \qquad 0 \leq t \leq 2\pi,$$

entonces la integral de línea $\int_\gamma F_1\,dx_1 + F_2\,dx_2$ es difícil de calcular directamente aunque se usen tablas. Sin embargo, podemos aplicar el teorema de Green a la región D entre γ y el círculo c de radio 1 y centro en el origen parametrizado por $(x_1, x_2) = (\cos t, \mathrm{sen}\,t)$, $0 \leq t \leq 2\pi$. Como la Ec. (4) es satisfecha, el teorema de Green implica que

$$\int_{\gamma \cup c^-} F_1\,dx_1 + F_2\,dx_2 = 0,$$

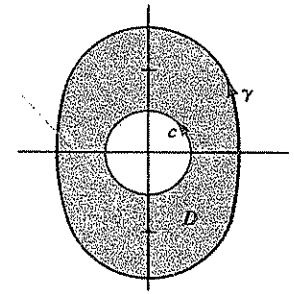


Figura 7

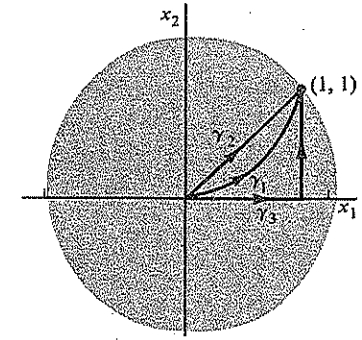


Figura 8

donde c^- es c recorrida en el sentido opuesto, es decir, con D a la derecha. Usando la Ec. 1.1, esta última ecuación puede ser escrita en la forma

$$\int_\gamma F_1\,dx_1 + F_2\,dx_2 = \int_c F_1\,dx_1 + F_2\,dx_2.$$

Pero en c tenemos que $x_1^2 + x_2^2 = 1$, de modo que

$$\begin{aligned}\int_\gamma F_1\,dx_1 + F_2\,dx_2 &= \int_c -x_2\,dx_1 + x_1\,dx_2 \\ &= \int_0^{2\pi} (\mathrm{sen}^2\,t + \cos^2 t)\,dt = 2\pi.\end{aligned}$$

Es importante notar que el teorema de Green no puede ser aplicado directamente a todo el interior de la elipse, ya que las derivadas $\partial F_2/\partial x_1$ y $\partial F_1/\partial x_2$ no existen en el origen.

Ejemplo 3. La curva γ_1 dada por $g(t) = (t, t^2)$, $0 \leq t \leq 1$, está dibujada en la Fig. 8. Supongamos que $F(x_1, x_2) = (F_1(x_1, x_2), F_2(x_1, x_2))$ es un campo vectorial diferenciable definido para $x_1^2 + x_2^2 < 4$ y que satisface la Ec. (4). Es posible que la integral de línea de F sobre γ_1 pueda ser calculada directamente en la forma

$$\int_{\gamma_1} F \cdot d\mathbf{x} = \int_0^1 [F_1(t, t^2) + F_2(t, t^2)(2t)]\,dt.$$

Sin embargo, hay otras posibilidades. Por ejemplo, la curva γ_2 puede ser parametrizada por $g_2(t) = (t, t)$, $0 \leq t \leq 1$. Como podemos aplicar el teorema de Green a la región entre γ_1 y γ_2, entonces como

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1\,dx_2 = 0$$

tenemos que

$$\int_{\gamma_1} F \cdot d\mathbf{x} + \int_{\gamma_2^-} F \cdot d\mathbf{x} = 0,$$

donde γ_2^- está dada por $g_2^-(t) = (1 - t, 1 - t)$ para $0 \le t \le 1$. Entonces, por la Ec. 1.1, las integrales sobre γ_1 y sobre γ_2 son iguales y esta última integral puede ser escrita en la forma

$$\int_{\gamma_2} F \cdot d\mathbf{x} = \int_0^1 [F_1(t, t) + F_2(t, t)]\, dt.$$

Otra alternativa es remplazar γ_1 por γ_3, donde γ_3 está parametrizada en dos pedazos por

$$g_3(t) = \begin{cases} \begin{pmatrix} t \\ 0 \end{pmatrix}, \\ \begin{pmatrix} 1 \\ t \end{pmatrix}, \end{cases} \qquad 0 \le t \le 1.$$

Por tanto,

$$\int_{\gamma_3} F \cdot d\mathbf{x} = \int_0^1 F_1(t, 0)\, dt + \int_0^1 F_2(1, t)\, dt.$$

Esto puede ser más fácil de calcular que las integrales sobre γ_1 y sobre γ_2, a pesar de que el resultado es siempre el mismo.

Las integrales de línea a lo largo de caminos *cerrados* a veces llamadas **circuitos**, son importantes y se distinguen de las otras usando el signo $\oint$. En el plano, esta notación tiene la ventaja adicional de que los símbolos $\oint$ y $\oint$ pueden ser usados para indicar que la curva se recorre en el sentido de los punteros del reloj en el segundo caso y en el sentido opuesto en el primero.

Ejemplo 4. El teorema de Green tiene dos aplicaciones físicas estrechamente relacionadas entre sí. Supongamos que D es una región en $\mathcal{R}^2$ cuya frontera es una curva simple γ orientada en el sentido contrario al de los punteros del reloj. Si γ tiene una parametrización lisa $g(t) = (g_1(t), g_2(t))$, $a \le t \le b$, con tangente distinta de cero en cada punto, entonces podemos formar el vector tangente unitario

$$\mathbf{t}(t) = \frac{g'(t)}{|g'(t)|} = \left(\frac{g_1'(t)}{|g'(t)|}, \frac{g_2'(t)}{|g'(t)|}\right)$$

y el vector normal unitario

$$\mathbf{n}(t) = \left(\frac{g_2'(t)}{|g'(t)|}, \frac{-g_1'(t)}{|g'(t)|}\right).$$

La Fig. 9 muestra un ejemplo. Si $F = (F_1, F_2)$ es un campo vectorial continuamente diferenciable definido en una región que contiene a D y a γ, en-

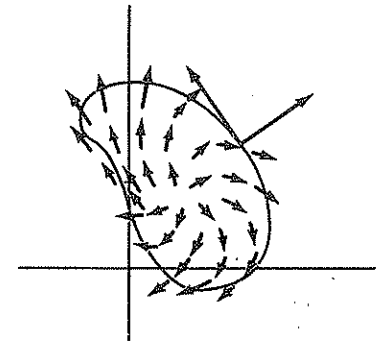


Figura 9

tonces la integral de línea del teorema de Green puede ser expresada en la forma

$$\oint_\gamma F_1\, dx_1 + F_2\, dx_2 = \int_a^b F(g(t)) \cdot \mathbf{t}(t)\, |g'(t)|\, dt$$
$$= \oint_\gamma F \cdot \mathbf{t}\, ds.$$

Definimos la función real rot F, llamada el **rotor** de F, por

$$\text{rot } F(\mathbf{x}) = \frac{\partial F_2}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_2}(\mathbf{x}).$$

y entonces el teorema de Green toma la forma

$$\int_D \text{rot } F\, dA = \oint_\gamma F \cdot \mathbf{t}\, ds,$$

que a veces se llama teorema de Stokes. Ahora, interpretamos F como un campo de fuerzas en el plano. La integral de línea representa el trabajo $W(\gamma)$ efectuado al mover una partícula a lo largo de γ. El teorema de Stokes dice que $W(\gamma)$ es igual a la integral sobre D del rotor de F. En particular, si rot F es idénticamente cero en D, entonces $W(\gamma) = 0$ para todo circuito liso contenido en D, sin importar la orientación de γ. Por supuesto, que es necesario que rot F esté definido en el interior del circuito γ en D para el cual se quiere deducir que $W(\gamma) = 0$. Conversamente, se puede demostrar que si $W(\gamma) = 0$, para todo circuito liso, entonces rot F es idénticamente cero (véase Sec. 2). F también puede ser interpretado como el campo de velocidad de flujo de un fluido en D. Es decir, el campo vectorial F representa, en cada punto D, la rapidez y dirección del flujo en ese punto. En este caso, la integral de línea que aparece en el teorema de Stokes se llama la **circulación** de F a lo largo de γ, y el teorema de Stokes dice que la circulación de F a lo largo de γ es la integral del rotor de F en D. Por tanto, decir que el rotor de F es idénticamente cero en D es lo mismo que

decir que la circulación a lo largo de todo circuito liso, cuyo interior está contenido en D, es cero. Por esta razón, un campo F para el cual rot F es cero, se llama **irrotacional**.

Ahora bien, si usamos el vector unitario normal, podemos reescribir el teorema de Green. En lugar de aplicar la Ec. (3) al campo (F_1, F_2), la aplicamos al par $(-F_2, F_1)$. Entonces la integral de línea es

$$\oint_\gamma -F_2\,dx_1 + F_1\,dx_2 = \int_a^b F(g(t))\cdot \mathbf{n}(t)\,|g'(t)|\,dt = \oint_\gamma F\cdot \mathbf{n}\,ds.$$

Por otra parte, la integral en D se transforma en

$$\int_D \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2}\right) dx_1\,dx_2.$$

Definimos la función real div F, llamada la **divergencia** de F por

$$\operatorname{div} F(\mathbf{x}) = \frac{\partial F_1}{\partial x_1}(\mathbf{x}) + \frac{\partial F_2}{\partial x_2}(\mathbf{x});$$

de modo que el teorema de Green toma la forma

$$\int_D \operatorname{div} F\,dA = \oint_\gamma F\cdot \mathbf{n}\,ds,$$

que, en esta forma, se llama el teorema de Gauss.

Usando la interpretación como flujo de un fluido, en la cual F es el campo velocidad de flujo, la integral de línea en el teorema de Gauss es la integral a lo largo de γ de la componente normal de F. Por esta razón, esta integral se llama el **flujo total** de F a través de γ en la dirección hacia afuera. El teorema de Gauss dice que el flujo $\Phi(\gamma)$ a través de γ es igual a la integral de divergencia de F en la región acotada por γ. Por tanto, div $F(x)$ mide la velocidad de cambio de la densidad de un flujo en el punto x. Si div $F(x)$ es predominantemente positivo en D, entonces el flujo hacia afuera $\Phi(\gamma)$ será positivo. Un flujo $\Phi(\gamma)$ es negativo, entonces más fluido entra en D de lo que sale. Si div F es idénticamente cero, entonces se dice que F representa un flujo **incompresible**.

EJERCICIOS

1. Use el teorema de Green para calcular el valor de la integral de línea $\int_\gamma x_2\,dx_1 + x_1^2\,dx_2$, donde γ es cada uno de los siguientes caminos cerrados:
   (a) El círculo definido por $g(t) = (\cos t, \operatorname{sen} t)$, $0 \le t \le 2\pi$.
   (b) El cuadrado con vértices en $(\pm 1, \pm 1)$, recorrido en el sentido opuesto al de las agujas del reloj.
   (c) El cuadrado con vértices en (0, 0), (1, 0), (1, 1) y (0, 1) recorrido en el sentido opuesto al de las agujas del reloj.

2. Calcule la integral de línea $\int_\gamma x_2\,dx_1 + x_1\,dx_2$, donde γ es una de las curvas del Ejemplo 3.
3. Sea γ la curva parametrizada por $g(t) = (2\cos t,\ 3\operatorname{sen} t)$, $0 \le t \le 2\pi$. Calcule $\int_\gamma (2x_1 + x_2)\,dx_1 + (x_1 + 3x_2)\,dx_2$.
4. Evalúe las siguientes integrales de línea usando el método que le parezca más sencillo.
   (a) $\int_\gamma e^x \cos y\,dx + e^x \operatorname{sen} y\,dy$, donde γ es el triángulo con vértices (0, 0), (1, 0), $(1, \pi/2)$ en el sentido opuesto al de las agujas del reloj.
   (b) Use el mismo integrando de la parte (a) y cambie el camino por el cuadrado con vértices en (0, 0), (1, 0), (1, 1) y (0, 1) recorrido en el sentido opuesto al de las agujas del reloj.
   (c) $\int_c (x^2 - y^2)\,dx + (x^2 + y^2)\,dy$, donde c es el círculo de radio 1 y centro (0, 0) recorrido en el sentido de las agujas del reloj.
5. Demuestre que si D es una región simple cuya frontera es una curva lisa por pedazos γ, entonces el área de D es
$$A(D) = \tfrac{1}{2}\oint_\gamma (-y\,dx + x\,dy).$$
6. Sea f una función real con segundas derivadas continuas en un subconjunto abierto D de $\mathcal{R}^2$. Sea F un campo vectorial definido en D por $F(\mathbf{x}) = \nabla f(\mathbf{x})$, el gradiente de f. Demuestre que si $F(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}))$, entonces la ecuación $\partial F_2/\partial x_1 - \partial F_1/\partial x_2 = 0$ se satisface en todo D.
7. Sea $f(x_1, x_2) = \operatorname{arctg}(x_2/x_1)$ para $x_1 > 0$. Suponga que el campo vectorial $\nabla f(x_1, x_2)$ tiene funciones coordenadas $F_1(x_1, x_2)$ y $F_2(x_1, x_2)$. Demuestre que F_1 y F_2 pueden ser extendidas de una manera natural a todo $(x_1, x_2) \neq (0, 0)$. Sea γ una curva lisa parametrizada de tal modo que vaya del punto (1, 0) al punto fijo (x_1, x_2), dando k vueltas en el sentido opuesto al de las agujas del reloj, alrededor del origen. Calcule la integral de línea
$$\int_\gamma \nabla f\cdot d\mathbf{x},$$
   (a) cuando $k = 0$, (b) cuando $k = 1$, (c) cuando k es un entero positivo arbitrario. (d) ¿Qué interpretación puede dársele a la integral cuando k es negativo?
8. Sea F un campo continuamente diferenciable definido en todos los puntos de $\mathcal{R}^2$ con la excepción de $\mathbf{x}_1$ y $\mathbf{x}_2$ y tal que satisface $\partial F_2/\partial x_1 - \partial F_1/\partial x_2 = 0$. Sean c_1 y c_2 dos círculos de centro $\mathbf{x}_1$ y $\mathbf{x}_2$ respectivamente y con radios menores que $|\mathbf{x}_1 - \mathbf{x}_2|$ y orientados en el sentido opuesto al de las agujas del reloj. Suponga que
$$\int_{ck} F\cdot d\mathbf{x} = I_k, \qquad k = 1, 2.$$
   Demuestre que si γ es un camino liso cerrado que no pasa por $\mathbf{x}_1$ ni por $\mathbf{x}_2$, entonces
$$\int_\gamma F\cdot d\mathbf{x} = n_1 I_1 + n_2 I_2,$$
   para ciertos enteros n_1 y n_2.
9. (a) Considere una partícula que se mueve en un plano vertical a la superficie de la tierra y que está sujeta al campo gravitacional $G(x, y) = (0, mg)$, donde m es

la masa de la partícula y g es la aceleración de gravedad. Demuestre que, al moverse la partícula en el plano, el trabajo efectuado es independiente del camino seguido, dependiendo solo de los puntos inicial y final. En particular, el trabajo efectuado cuando la partícula se mueve en un camino cerrado es cero.

(b) Remplace el campo G por un campo $F = (F_1, F_2)$ que satisface $\partial F_2/\partial x_1 = \partial F_1/\partial x_2$ en todo el plano. Demuestre que las mismas conclusiones son verdaderas.

10. (a) Sea $\mathfrak{R} \xrightarrow{g} \mathfrak{R}^2$ una parametrización de una curva cerrada simple γ recorrida en el sentido de las agujas del reloj. Demuestre que si el vector unitario tangente a γ es $\mathbf{t}(t) = g'(t)/|g'(t)|$, entonces el vector unitario normal a γ dirigido *hacia afuera* está dado por $\mathbf{n}(t) = (g_2'(t)/|g'(t)|, -g_1'(t)/|g'(t)|)$, donde g_1 y g_2 son las funciones coordenadas de g.

(b) Demuestre que el teorema de Green puede ser escrito en la forma

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1\, dx_2 = \int_\gamma F \cdot \mathbf{t}\, ds,$$

donde $F = (F_1, F_2)$ y ds denota la integración con respecto del largo del arco.

(c) Demuestre que el teorema de Green también puede ser expresado en la forma

$$\int_D \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2}\right) dx_1\, dx_2 = \int_\gamma F \cdot \mathbf{n}\, ds.$$

[*Indicación*: Remplace F_2 por F_1 y F_1 por $-F_2$ en la fórmula anterior.]

11. Suponga que el campo vectorial $F = (F_1, F_2)$ del Ejercicio 10(c) es un campo gradiente, es decir, $F = \nabla f$ para cierta función real f. Demuestre que el Teorema de Green puede ser expresado en la forma

$$\int_D \Delta f\, dA = \int_\gamma \nabla f \cdot \mathbf{n}\, ds,$$

donde $\Delta f = \partial^2 f/\partial x_1^2 + \partial^2 f/\partial x_2^2$ es el laplaciano de f.

12. (a) Demuestre que si f es una función real continua definida en un subconjunto abierto D de $\mathfrak{R}^2$ y $\int_M f\, dA = 0$ para todo disco circular M en D, entonces f es idénticamente cero en D. [*Indicación*: Demuestre que si $f(\mathbf{x}_0) \neq 0$ para algún $\mathbf{x}_0$ en D entonces existe un disco M de centro $\mathbf{x}_0$ tal que $|f(\mathbf{x})| \geq \delta$ para algún $\delta > 0$ y todo $\mathbf{x}$ en M.]

(b) Use la parte (a) y el teorema de Stokes para demostrar que si rot F es continua en un conjunto abierto D, y la circulación de F es cero alrededor de todo circuito liso en D, entonces F es irrotacional en D, es decir, rot F es idénticamente cero en D.

(c) Use la parte (a) y el teorema de Gauss para demostrar que si div F es continua en D y $\Phi(\gamma) = 0$ para todo circuito liso γ en D, entonces F es incompresible.

13. Las ecuaciones rot $F = 0$ y div $F = 0$ aparecen en la teoría de variables complejas en una forma ligeramente diferente como las ecuaciones de Riemann. Demuestre que si $u(x, y)$ y $v(x, y)$ son la parte real y la parte imaginaria respectivamente, de las siguientes funciones complejas, entonces el campo vectorial $F(x, y) = (v(x, y), u(x, y)$ es irrotacional e incompresible.

(a) $(x + iy)^2$.

(b) e^{x+iy}.

(c) $(\frac{1}{2}) \log (x^2 + y^2) + i \operatorname{arctg}\, y/x$, $x > 0$.

2. Campos vectoriales conservativos

Los ejemplos de la sección anterior muestran que, bajo ciertas condiciones, es posible alterar el camino de integración de una integral de línea sin que varíe su valor. No todas las integrales de línea tienen esta propiedad, pero las que la tienen son especialmente importantes, no solo por las razones computacionales que ya han sido ilustradas, sino también por su relación con el gradiente. Tenemos, en efecto, el siguiente teorema que también es válido en $\mathfrak{R}^n$ y que es el converso del Teorema 2.5 del Cap. 4.

2.1 Teorema

Sea F un campo vectorial continuo definido en un subconjunto abierto poligonalmente conexo D de $\mathfrak{R}^n$. Si la integral de línea

$$\int_\gamma F \cdot d\mathbf{x}$$

es independiente del camino liso por pedazos γ de $\mathbf{x}_0$ a $\mathbf{x}$ en D, entonces la función real definida por

$$f(\mathbf{x}) = \int_{\mathbf{x}_0}^{\mathbf{x}} F \cdot d\mathbf{x}$$

es continuamente diferenciable y satisface la ecuación vectorial $\nabla f = F$ en todo D.

Demostración. Debemos demostrar que, para todo $\mathbf{x}$ en D, $\nabla f(\mathbf{x}) = F(\mathbf{x})$. Como $\mathbf{x}$ es un punto interior de D, existe una bola de radio δ y centro $\mathbf{x}$ contenida en D. Esto implica que para todo vector unitario $\mathbf{u}$ y todo real t tal que $|t| < \delta$, los vectores $\mathbf{x} + t\mathbf{u}$ están en D. Como la integral de línea es independiente del camino, elegimos un camino en D liso por pedazos, arbitrario y que une los puntos $\mathbf{x}_0$ y $\mathbf{x}$. Extendemos este camino por un segmento de recta a los vectores $\mathbf{x} + t\mathbf{u}$, $|t| < \delta$, tal como lo muestra la Fig. 10. Para demostrar que f es continuamente diferenciable, notamos que

$$\begin{aligned} f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x}) &= \int_{\mathbf{x}_0}^{\mathbf{x}+t\mathbf{u}} F \cdot d\mathbf{x} - \int_{\mathbf{x}_0}^{\mathbf{x}} F \cdot d\mathbf{x} \\ &= \int_{\mathbf{x}}^{\mathbf{x}+t\mathbf{u}} F \cdot d\mathbf{x} \\ &= \int_0^t F(\mathbf{x} + v\mathbf{u}) \cdot \mathbf{u}\, dv. \end{aligned}$$

Figura 10

Entonces, haciendo $\mathbf{u} = \mathbf{e}_j$, el j-ésimo vector de la base natural de $\mathcal{R}^n$, tenemos que

$$\begin{aligned}\frac{\partial f}{\partial x_j}(\mathbf{x}) &= \lim_{t\to 0}\frac{f(\mathbf{x}+t\mathbf{e}_j)-f(\mathbf{x})}{t}\\ &= \lim_{t\to 0}\frac{1}{t}\int_0^t F(\mathbf{x}+v\mathbf{e}_j)\cdot\mathbf{e}_j\,dv\\ &= \frac{d}{dt}\int_0^t F(\mathbf{x}+v\mathbf{e}_j)\cdot\mathbf{e}_j\,dv\Big|_{t=0}\\ &= F(\mathbf{x})\cdot\mathbf{e}_j = F_j(\mathbf{x}),\end{aligned}$$

donde F_j es la j-ésima función coordenada de F. Como supusimos que F es continua, también lo son las derivadas parciales $\partial f/\partial x_j$, de modo que f es continuamente diferenciable en D. Por último, las ecuaciones $\partial f/dx_j(\mathbf{x}) = F_j(\mathbf{x})$, $j = 1, \ldots, n$, indican que $\nabla f = F$ en D.

Un campo vectorial F para el cual existe una función real f tal que $\nabla f = F$, se llama **campo conservativo** o **gradiente**. En ese caso, f se llama el **potencial** de F. La razón para esta terminología se indica en el ejemplo que sigue.

Ejemplo 1. Supongamos que un campo continuo de fuerzas F, definido en una región D en $\mathcal{R}^3$, es tal que el trabajo efectuado al mover una partícula de un punto a otro bajo la influencia de este campo, es independiente del camino usado para llegar de un punto al otro. Así, si $\mathbf{x}_1$ y $\mathbf{x}_2$ son dos puntos en el campo y $W(\mathbf{x}_1, \mathbf{x}_2)$ es el trabajo efectuado al ir de $\mathbf{x}_1$ a $\mathbf{x}_2$, entonces

$$W(\mathbf{x}_1, \mathbf{x}_2) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} F\cdot d\mathbf{x}.$$

Si la partícula sigue el camino definido por $g(t)$, entonces los vectores velocidad y aceleración son $\mathbf{v}(t) = g'(t)$ y $\mathbf{a}(t) = g''(t)$ y tenemos que $F(g(t)) = m\mathbf{a}(t)$, donde m es la masa de la partícula. Entonces

$$W(\mathbf{x}_1, \mathbf{x}_2) = \int_{t_1}^{t_2} m\mathbf{a}(t)\cdot\mathbf{v}(t)\,dt,$$

si $g(t_i) = \mathbf{x}_i$. Pero como $\mathbf{a}(t) = \mathbf{v}'(t)$ y $(d/dt)v^2(t) = 2\mathbf{v}(t)\cdot\mathbf{v}'(t)$, tenemos que

$$\begin{aligned}W(\mathbf{x}_1, \mathbf{x}_2) &= \frac{m}{2}\int_{t_1}^{t_2}\frac{d}{dt}\,[v^2(t)]\,dt,\\ &= \frac{m}{2}\left(v^2(t_2) - v^2(t_a)\right).\end{aligned} \tag{1}$$

La función $k(t) = (m/2)v^2(t)$ se llama la **energía cinética** de la partícula en el instante t.

Por otra parte, si fijamos un punto $\mathbf{x}_0$ en D, entonces por el Teorema 2.1, la ecuación

$$u(\mathbf{x}) = -\int_{\mathbf{x}_0}^{\mathbf{x}} F\cdot d\mathbf{x}$$

define una función continuamente diferenciable u en D. Usando la independencia del camino para integrar de $\mathbf{x}_1$ a $\mathbf{x}_2$ pasando por $\mathbf{x}_0$, obtenemos que

$$\begin{aligned}W(\mathbf{x}_1, \mathbf{x}_2) &= \int_{\mathbf{x}_1}^{\mathbf{x}_2} F\cdot d\mathbf{x}\\ &= \int_{\mathbf{x}_0}^{\mathbf{x}_2} F\cdot d\mathbf{x} - \int_{\mathbf{x}_0}^{\mathbf{x}_1} F\cdot d\mathbf{x}\\ &= -u(\mathbf{x}_2) + u(\mathbf{x}_1).\end{aligned} \tag{2}$$

Si comparamos las Ecs. (1) y (2), deducimos que

$$u(\mathbf{x}_2) + \frac{m}{2}v^2(t_2) = u(\mathbf{x}_1) + \frac{m}{2}v^2(t_1).$$

En otras palabras, la suma $u(g(t)) + k(t)$ es una constante independiente de t a lo largo de la curva definida por g y se llama la energía total del camino. Por esta razón, la función $u(\mathbf{x})$, que depende solo de la posición en D, se llama la **energía potencial** del campo F. Nótese que hay una elección arbitraria al definir el potencial, ya que se ha considerado que el punto $\mathbf{x}_0$ tiene potencial cero. Si cambiamos $\mathbf{x}_0$ por $\mathbf{x}_1$ la función u se altera por la suma de una constante igual a $W(\mathbf{x}_0, \mathbf{x}_1)$. La energía total es siempre constante, es decir, se «conserva», razón por la cual se habla de «campos conservativos».

Si F es un campo vectorial definido en una región D de $\mathcal{R}^n$, la independencia del camino de la integral $\int_\gamma F\cdot d\mathbf{x}$ significa, por definición, que

$$\int_{\gamma[\mathbf{x}_1\mathbf{x}_2]} F\cdot d\mathbf{x} = \int_{\delta[\mathbf{x}_1\mathbf{x}_2]} F\cdot d\mathbf{x} \tag{2.2}$$

donde $\gamma[\mathbf{x}_1, \mathbf{x}_2]$ y $\delta[\mathbf{x}_1, \mathbf{x}_2]$ son dos curvas lisas por pedazos en D, con punto inicial $\mathbf{x}_1$ y punto final $\mathbf{x}_2$. Otra manera de formular esta propiedad es

$$\int_\gamma F\cdot d\mathbf{x} = 0 \tag{2.3}$$

para toda curva cerrada lisa por pedazos, γ en D. La equivalencia de estas dos propiedades se obtiene notando que $\gamma[\mathbf{x}_1, \mathbf{x}_2]$ seguido de $\delta[\mathbf{x}_1, \mathbf{x}_2]$ en dirección contraria, es un camino cerrado y que un camino cerrado puede ser separado en puntos $\mathbf{x}_1$ y $\mathbf{x}_2$ en dos caminos distintos que unen a $\mathbf{x}_1$ y $\mathbf{x}_2$. Los detalles de la demostración han sido ilustrados en la Sec. 1 y serán dejados como ejercicio.

Resumimos lo que hemos demostrado acerca de campos gradientes en el Teorema 2.5 del Cap. 4 y el Teorema 2.1 de esta sección y las observaciones anteriores en el siguiente teorema.

2.4 Teorema

Sea F un campo vectorial continuo definido en un conjunto abierto y poligonalmente conexo D en $\mathcal{R}^n$. Entonces las afirmaciones siguientes son equivalentes:

1. F es el gradiente de una función continuamente diferenciable f en D.
2. La integral de línea de F a lo largo de un camino que va de $\mathbf{x}_1$ a $\mathbf{x}_2$ es independiente de la curva lisa por pedazos γ entre $\mathbf{x}_1$ y $\mathbf{x}_2$, de modo que puede ser escrita en la forma

$$\int_{\mathbf{x}_1}^{\mathbf{x}_2} F \cdot d\mathbf{x}.$$

3. Para toda curva cerrada y lisa por pedazos en D se tiene que

$$\oint F \cdot d\mathbf{x} = 0.$$

Un criterio más intrínseco para decidir si un campo vectorial continuo es o no un campo gradiente se obtiene como sigue: Supongamos primero que $\mathcal{R}^2 \xrightarrow{F} \mathcal{R}^2$ es continuo en un conjunto abierto D y que F es un campo gradiente, es decir, que existe una función real f definida en D tal que $\nabla f = F$. En términos de las funciones coordenadas F_1 y F_2 de F, esto significa que

$$\frac{\partial f}{\partial x_1} = F_1 \quad \text{y} \quad \frac{\partial f}{\partial x_2} = F_2.$$

Si F es continuamente diferenciable, podemos considerar las segundas derivadas parciales

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial F_1}{\partial x_2} \quad \text{y} \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial F_2}{\partial x_1},$$

y concluir que

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1} \tag{3}$$

en todo D. Por definición de rot F, la Ec. (3) dice que rot $F = 0$. Esta ecuación puede ser expresada de otra manera. Consideraremos la situación más general de un campo $\mathcal{R}^n \xrightarrow{F} \mathcal{R}^n$ continuamente diferenciable en un subconjunto abierto D de $\mathcal{R}^n$. Si F es un campo gradiente, entonces existe f tal que $\nabla f = F$ o, en términos de funciones coordenadas,

$$\frac{\partial f}{\partial x_j} = F_j, \qquad j = 1, \ldots, n.$$

Derivando con respecto de x_i, obtenemos

$$\frac{\partial F_j}{\partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x} = \frac{\partial F_i}{\partial x_j}. \tag{4}$$

Pero las funciones $\partial F_i/\partial x_j$ son los elementos de la matriz jacobiana de $\mathcal{R}^n \xrightarrow{F} \mathcal{R}^n$ y la Ec. (4) dice que esta matriz es simétrica. Por tanto,

2.5 Teorema

Si $\mathcal{R}^n \xrightarrow{F} \mathcal{R}^n$ es un campo gradiente continuamente diferenciable, entonces la matriz jacobiana de F es simétrica.

Ejemplo 2. La conversa del Teorema 2.5 es falsa, como se ve considerando el campo

$$F(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right),$$

definido para todo $(x, y) \neq 0$. Es fácil verificar que $\partial F_1/\partial y = \partial F_2/\partial x$. Sin embargo, no existe f tal que $\nabla f = F$. La razón para esto es que, para $x > 0$, la función $f(x, y) = \operatorname{arctg}(y/x)$ satisface $\nabla f = F$, pero esta f no puede ser extendida a una solución de la ecuación en el plano sin el origen.

El ejemplo 2 muestra que la naturaleza de la región D en la cual F está definida es importante para determinar si F es un campo gradiente. Haciendo una suposición especial para D, podemos obtener un converso parcial del Teorema 2.5.

2.6 Teorema

Sea R un rectángulo coordenado abierto en $\mathcal{R}^n$ y sea F un campo continuamente diferenciable en R. Si la matriz jacobiana $F'(\mathbf{x})$ es simétrica en R, entonces F es un campo gradiente.

Demostración. Elijamos un punto fijo $\mathbf{x}_0$ en R y sea $\mathbf{x}$ cualquier punto de R. Consideramos caminos de $\mathbf{x}_0$ a $\mathbf{x}$ formados por una sucesión de segmentos de recta paralelos a los ejes y tales que cada variable varía a lo más en uno de los segmentos. La Fig. 11 muestra algunos ejemplos en tres dimensiones. La razón

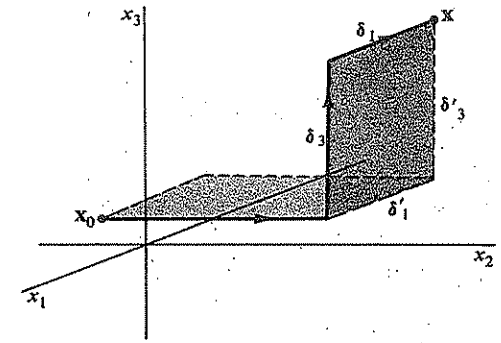


Figura 11

para considerar este tipo de caminos es que con ellos nos podemos acercar a x en cualquier dirección coordenada, lo cual es importante al tomar las derivadas parciales en x. Elegimos uno de estos caminos, lo llamamos $\gamma_{\mathbf{x}}$ y definimos una función real f por

$$f(\mathbf{x}) = \int_{\gamma_{\mathbf{x}}} F \cdot d\mathbf{x}. \tag{5}$$

A pesar de que $\gamma_{\mathbf{x}}$ ha sido elegida entre varios caminos posibles, veremos que cada una de las elecciones posibles conduce al mismo valor para $f(\mathbf{x})$. La razón es que cada uno de estos caminos puede ser alterado paso a paso en cualquier otro haciendo cambios que dejan inalterado el valor de la integral (5). Cada camino puede ser descrito como una sucesión de direcciones coordenadas, de las cuales solo una se deja variar cada vez. (Por ejemplo, el camino punteado de la Fig. 11 corresponde a x_1, x_2, x_3 y el camino sólido a x_2, x_3, x_1.) Es claro que el paso de una de estas sucesiones a la otra puede ser hecha intercambiando pares de variables adyacentes hasta obtener el orden deseado. Pero el intercambio remplaza un par de segmentos (δ_i, δ_j) por otro par (δ_i', δ_j') por otro en el mismo plano bidimensional. Para comprobar que este remplazo deja invariante el valor de la integral, formamos el circuito δ formado por los segmentos δ_i y δ_j, seguidos por δ_i' y δ_j' sus direcciones invertidas. En estos segmentos, x_i y x_j son las únicas variables que cambian, de modo que la integral en el circuito se puede escribir en la forma

$$\oint_\delta F \cdot d\mathbf{x} = \oint_\delta F_i\, dx_j + F_j\, dx_i.$$

Aplicamos el teorema de Green al rectángulo bidimensional R_δ cuya frontera es δ, obteniendo

$$\oint_\delta F \cdot d\mathbf{x} = \int_{R_\delta} \left(\frac{\partial F_i}{\partial x_i} - \frac{\partial F_i}{\partial x_j} \right) dx_i\, dx_j = 0,$$

ya que, por la hipótesis de simetría, $\partial F_j/\partial x_i - \partial F_i/\partial x_j = 0$. Por tanto,

$$\oint_\delta F \cdot d\mathbf{x} = \int_{(\delta_i, \delta_j)} F \cdot d\mathbf{x} - \int_{(\delta_j', \delta_i')} F \cdot d\mathbf{x} = 0,$$

de modo que el cambio de camino deja invariante el valor de la integral.

Una vez que se ha establecido que x puede ser aproximado a lo largo de un camino de integración que varía solo en una coordenada arbitraria, digamos la k-ésima, obtenemos, como en la demostración del Teorema 2.1, la ecuación $\partial f/\partial x_k(\mathbf{x}) = F_k(\mathbf{x})$ para todo k. Por tanto, $\nabla f(\mathbf{x}) = F(\mathbf{x})$ para todo x en R.

Ejemplo 3. Aplicando el teorema al campo

$$F(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right), \qquad (x, y) \neq (0, 0),$$

del Ejemplo 2, concluimos que F restringido a cualquier rectángulo coordenado que no contiene al origen es un campo gradiente. Por ejemplo, esto sucede en cada uno de los cuatro semiplanos acotados por un eje coordenado. Una función potencial f para el semiplano $x > 0$ puede ser calculado por la integral de línea

$$f(x, y) = \int_{(1,0)}^{(x,y)} \frac{-y\, dx}{x^2 + y^2} + \frac{-x\, dy}{x^2 + y},$$

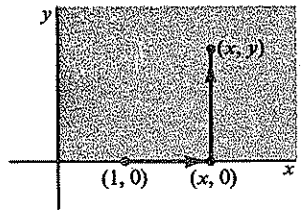


Figura 12

donde el camino de integración es cualquier curva lisa por pedazos desde (1, 0) a (x, y). Un camino particularmente sencillo es el formado por un segmento de recta de (1, 0) a $(x, 0)$ seguido por otro de $(x, 0)$ a (x, y). En el primer segmento la integral es cero, ya que y es idénticamente cero y en el segundo segmento, como x es constante, la integral se reduce a

$$\int_0^y \frac{x\, dy}{x^2 + y^2} = \operatorname{arctg}\left(\frac{y}{x}\right).$$

El potencial más general para F en el semiplano derecho difiere de éste en, a lo más, una constante. (¿Por qué?) De modo que allí la solución general de $\nabla f = F$ es

$$f(x, y) = \operatorname{arctg} \frac{y}{x} + c.$$

Compare el método de solución descrito en el Ejercicio 11 de la Sec. 2 del Cap. 4.

EJERCICIOS

1. Considere la aproximación del campo gravitacional de la tierra cuando actúa sobre una partícula de masa 1 representada por el campo vectorial $F(x, y, z) = (0, 0, -g)$.
   (a) Encuentre la función potencial $u(x, y, z)$ para F que es cero cuando $(x, y, z) = (0, 0, 0)$.
   (b) Si una partícula de masa 1 tiene en (0, 0, 0) un vector velocidad (v_1, v_2, v_3) con $v_3 > 0$ y la única fuerza que actúa sobre la partícula es F, encuentre el camino que sigue la partícula.
   (c) Verifique que la suma de la energía potencial y la energía cinética se mantiene constante a lo largo de la curva de la parte (b).

2. (a) Demuestre que si F y G son campos gradientes definidos en el mismo dominio entonces $F + G$ y cF son campos gradientes donde c es una constante.
   (b) Sea $\mathcal{U}$ el espacio vectorial de los campos gradientes definidos en un dominio D. Demuestre que $\mathcal{U}$ tiene dimensión infinita.

3. Use los Teoremas 2.4 o 2.6 para decidir si los siguientes campos vectoriales son campos gradientes.
   (a) $F(x, y) = (x - y, x + y)$, para (x, y) en $\mathcal{R}^2$.
   (b) $G(x, y, z) = (y, z, x)$, para (x, y, z) en $\mathcal{R}^3$.
   (c) $H(x, y) = \left(\dfrac{-y}{x^2 + y^2}, \dfrac{x}{x^2 + y^2}\right)$, $(x, y) \neq (0, 0)$.

4. Use el Teorema 2.5 para demostrar que los campos vectoriales de los Ejercicios 3(a) y 3(b) no son campos gradientes en ningún subconjunto abierto de $\mathcal{R}^2$ y $\mathcal{R}^3$ respectivamente.

5. Demuestre que el campo vectorial del Ejercicio 3(c) es un campo gradiente en la región $y > 0$ de $\mathcal{R}^2$ y encuentre una representación explícita para su potencial.

6. Considere el campo vectorial definido para todo (x, y, z) distinto de $(0, 0, z)$ por

$$F(x, y, z) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0\right).$$

   ¿Es F un campo gradiente?

7. Encuentre un potencial para cada uno de los campos siguientes:
   (a) $F(x, y, z) = (2xy, x^2 + z^2, 2yz)$.
   (b) $G(x, y) = (y \cos xy, x \cos xy)$.
   (c) $L(x_1, x_2) = \left(\dfrac{x_1}{x_1^2 + x_2^2}, \dfrac{x_2}{x_1^2 + x_2^2}\right)$, con $(x_1, x_2) \neq (0, 0)$.

8. Considere el campo vectorial F que es el gradiente del potencial newtoniano $f(\mathbf{x}) = -|\mathbf{x}|^{-1}$ para $\mathbf{x}$ distinto de cero en $\mathcal{R}^3$. Encuentre el trabajo efectuado al mover una partícula de $(1, 1, 1)$ a $(-2, -2, -2)$ a lo largo de una curva lisa en el dominio de F.

9. Dé una demostración detallada de la equivalencia entre las relaciones 2.2 y 2.3.

10. ¿Cuántos caminos del tipo descrito en la demostración del Teorema 2.6 existen entre los puntos $\mathbf{x}_0$ y $\mathbf{x}$ de $\mathcal{R}^n$?

3. Integrales de superficie

En la Sec. 4 del Cap. 3 hemos definido la integral de una función real y de un campo vectorial a lo largo de una curva lisa. Definir una integral sobre una superficie S conduce a una situación geométrica distinta, pero que es muy análoga a la integral de línea. Para empezar, suponemos que S está parametrizada por una función continuamente diferenciable $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$. Escribiremos g en la forma

$$g(u, v) = \begin{pmatrix} g_1(u, v) \\ g_2(u, v) \\ g_3(u, v) \end{pmatrix}, \tag{1}$$

donde $\mathbf{u} = (u, v)$ está en un conjunto D en $\mathcal{R}^2$ que suponemos tiene como frontera una curva lisa por pedazos. Suponemos además, que en cada punto $g(u, v)$ de S los vectores tangentes definidos por las derivadas parciales vectoriales

$$\frac{\partial g}{\partial u}(u, v), \qquad \frac{\partial g}{\partial v}(u, v)$$

determinan un plano tangente a S, en otras palabras, suponemos que las dos tangentes son linealmente independientes. Si S satisface todas las condiciones indicadas, diremos que S es un **pedazo de superficie lisa.**

En una curva lisa, una orientación queda determinada por la elección de una parametrización que va desde un extremo al otro. Análogamente, en un pedazo de superficie lisa, una parametrización inyectiva determina un vector normal especial:

$$\frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \tag{2}$$

en $g(u, v)$. Véase la Fig. 13.

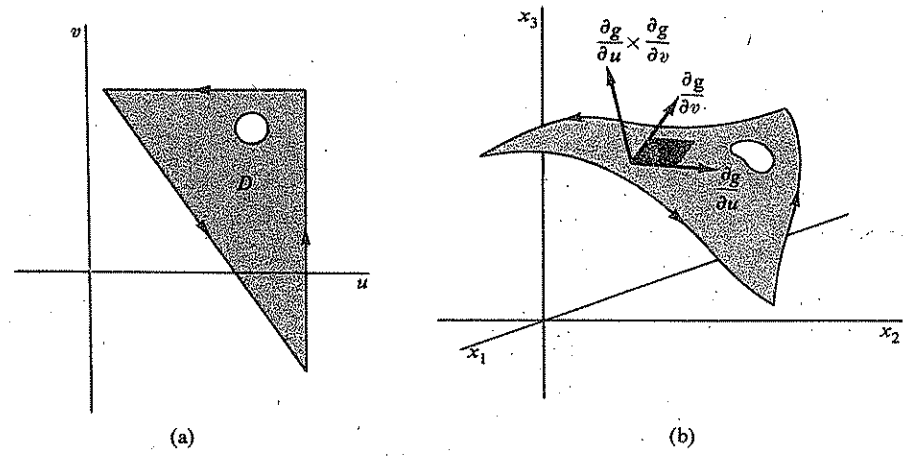


Figura 13

Recordemos que el largo del producto cruz de dos vectores a y b es el área del paralelogramo generado por a y b. En particular,

$$\left|\frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v)\right|$$

representa el área del paralelogramo tangente mostrado en la Fig. 13(b). Si consideramos que la construcción de tales paralelogramos en los puntos $g(u_k, v_k)$

correspondientes a los puntos (u_k, v_k) de un reticulado en D, entonces es natural definir el **área** de S por

$$\sigma(S) = \int_D \left| \frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \right| du\, dv. \tag{3.1}$$

Supongamos que g es inyectiva. La integral sobre D existe, ya que se supuso que g es continuamente diferenciable en D. Análogamente, si p es una función real continua definida en S, definimos la integral de p en S por

$$\int_S p\, d\sigma = \int_D p(g(u, v)) \left| \frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \right| du\, dv. \tag{3.2}$$

Esta definición es análoga a la de la Ec. (2) de la Sec. 4 del Cap. 3 para la integral de una función real a lo largo de una curva lisa.

Ejemplo 1. Sea S parametrizada por

$$g(u, v) = \begin{pmatrix} u \\ v \\ u^2 + v^2 \end{pmatrix},$$

para

$$1 \leq u^2 + v^2 \leq 4;$$

de modo que S es el gráfico de $x^2 + y^2$ para $1 \leq x^2 + y^2 \leq 4$. Esta superficie está representada en la Fig. 14. Entonces

$$\frac{\partial g}{\partial u}(u, v) = \begin{pmatrix} 1 \\ 0 \\ 2u \end{pmatrix}, \qquad \frac{\partial g}{\partial v}(u, v) = \begin{pmatrix} 0 \\ 1 \\ 2v \end{pmatrix}.$$

Tenemos que

$$\frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) = \left(\begin{vmatrix} 0 & 1 \\ 2u & 2v \end{vmatrix}, \begin{vmatrix} 2u & 2v \\ 1 & 0 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \right).$$

Por tanto,

$$\sigma(S) = \int_{1 \leq u^2+v^2 \leq 4} \sqrt{4u^2 + 4v^2 + 1}\, du\, dv$$

$$= \int_0^{2\pi} d\theta \int_1^2 \sqrt{4r^2 + 1}\, r\, dr = \frac{\pi}{6}[17^{3/2} - 5^{3/2}].$$

Si S tiene una distribución no uniforme de densidad igual en cada punto a la distancia del punto al eje de simetría de S, entonces la densidad puede ser representada por la función $p(x, y, z) = \sqrt{x^2 + y^2}$. En un punto $g(u, v)$ de S, la

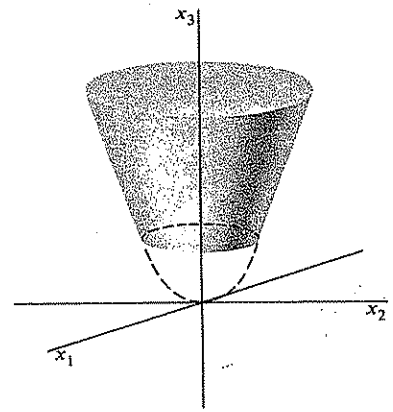


Figura 14

densidad será $p(g(u, v)) = \sqrt{u^2 + v^2}$. Podemos interpretar la integral de p sobre S como la masa total de la superficie y tenemos que

$$\int_S p\, d\sigma = \int_{1 \leq u^2+v^2 \leq 4} \sqrt{u^2 + v^2}\sqrt{4u^2 + 4v^2 + 1}\, du\, dv$$

$$= \int_0^{2\pi} d\theta \int_1^2 r^2\sqrt{4r^2 + 1}\, dr.$$

El propósito principal de esta sección es la definición de la integral de un campo vectorial continuo $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}$ sobre una superficie S. Bajo la suposición de que S es un pedazo de superficie lisa representado por la función de la Ec. (1), comparamos el vector normal $\partial g/\partial u \times \partial g/\partial v$ con el campo vectorial en el punto $g(u, v)$ de S. La Fig. 15 muestra estos vectores en un punto. Si $\mathbf{n}$ es un vector *unitario* normal a S en $g(u, v)$, entonces el producto punto $F \cdot \mathbf{n}$ en $g(u, v)$ es la coordenada de F en la dirección de $\mathbf{n}$. Pero como

$$\mathbf{n} = \frac{\dfrac{\partial g}{\partial u} \times \dfrac{\partial g}{\partial v}}{\left| \dfrac{\partial g}{\partial u} \times \dfrac{\partial g}{\partial v} \right|},$$

tenemos que

$$F(g(u, v)) \cdot \left(\frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \right)$$

es igual a la coordenada de $F(g(u, v))$ en la dirección de $\mathbf{n}$, multiplicada por el área del paralelogramo tangente generado por $\partial g/\partial u$ y $\partial g/\partial v$ en $g(u, v)$. Definimos la **integral de superficie** de F en S por

$$\int_D F(g(u, v)) \cdot \left(\frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \right) du\, dv, \tag{3.3}$$

y la denotamos por $\int_S F \cdot dS$ o por $\int_S F \cdot \mathbf{n}\, d\sigma$

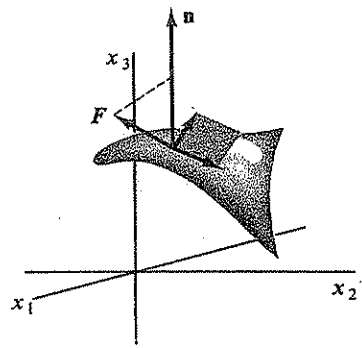


Figura 15

Es fácil verificar que las coordenadas del vector normal están relacionadas con los determinantes jacobianos por

$$\frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v} = \left(\frac{\partial(x_2, x_3)}{\partial(u, v)}, \frac{\partial(x_3, x_1)}{\partial(u, v)}, \frac{\partial(x_1, x_2)}{\partial(u, v)} \right),$$

donde x_1, x_2 y x_3 representan las funciones coordenadas de g. Si el campo vectorial F tiene funciones coordenadas F_1, F_2, F_3, entonces la integral de superficie se escribe a menudo en la forma

$$\int_S F \cdot dS = \int_D F_1 \frac{\partial(x_2, x_3)}{\partial(u, v)} + F_2 \frac{\partial(x_3, x_1)}{\partial(u, v)} + F_3 \frac{\partial(x_1, x_2)}{\partial(u, v)} \, du \, dv$$

$$= \int_D F_1 \, dx_2 \, dx_3 + F_2 \, dx_3 \, dx_1 + F_3 \, dx_1 \, dx_2.$$

La última abreviación es especialmente conveniente y su significado será discutido en la Sec. 7.

Ejemplo 2. Supongamos que el campo vectorial continuo $\mathbb{R}^3 \xrightarrow{F} \mathbb{R}^3$ describe la rapidez y dirección de flujo de un fluido en cada punto de una región R en la cual F está definido. Usando la integral de superficie definiremos el flujo por unidad de área y tiempo a través de un pedazo de superficie lisa en en R. Si S es plano y F es un campo constante, entonces el flujo es igual a $F_n \sigma(S)$, donde F_n es la coordenada de F en la dirección de un vector unitario normal a S. Por tanto, en este caso, el flujo es igual al volumen de un tubo de fluido como el ilustrado en la Fig. 16. Como $F_n = F \cdot \mathbf{n}$, tenemos que el flujo es

$$\Phi = F \cdot \mathbf{n} \sigma(S)$$

Si S es un pedazo de superficie lisa en R, dividimos S a lo largo de curvas coordenadas de la forma $u =$ constante y $v =$ constante y suponemos que el

campo F es constante en cada uno de los pedazos de S así formados. Aproximando S por los paralelogramos generados por los vectores

$$\Delta u \frac{\partial g}{\partial u} \quad \text{y} \quad \Delta v \frac{\partial g}{\partial v}$$

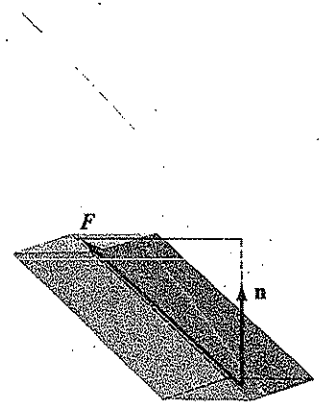


Figura 16

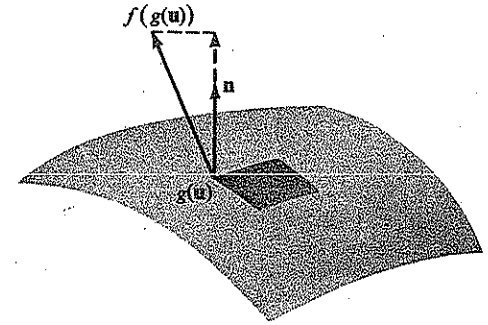


Figura 17

tenemos una situación como la de la Fig. 17. El flujo aproximado a través de una subdivisión S_k de S tiene la forma

$$\Phi_k = F(g(\mathbf{u}_k)) \cdot \mathbf{n}_k \sigma(S_k)$$

$$= F(g(\mathbf{u}_k)) \cdot \left(\frac{\partial g}{\partial u}(\mathbf{u}_k) \times \frac{\partial g}{\partial v}(\mathbf{u}_k) \right) \Delta u \, \Delta v.$$

La suma

$$\sum_{k=1}^{N} \Phi_k = \sum_{k=1}^{N} F(g(\mathbf{u}_k)) \cdot \left(\frac{\partial g}{\partial u}(\mathbf{u}_k) \times \frac{\partial g}{\partial v}(\mathbf{u}_k) \right) \Delta u \, \Delta v$$

representa una aproximación de lo que quisiéramos llamar el flujo a través de S. Esta aproximación será tanto mejor cuanto más fina sea la subdivisión de S. Pero esta subdivisión es tanto más fina cuanto más fino es el reticulado correspondiente G en el dominio D de parámetros. Por otra parte, si F es continuo en S y g es continuamente diferenciable en D, entonces

$$\lim_{m(G) \to 0} \sum_{k=1}^{N} \Phi_k = \int_D F(g(\mathbf{u})) \cdot \left(\frac{\partial g}{\partial u}(\mathbf{u}) \times \frac{\partial g}{\partial v}(\mathbf{u}) \right) d\sigma$$

$$= \int_D F \cdot d\omega,$$

que es la integral de F sobre S, como la definiéramos anteriormente. Por tanto, definimos el **flujo** de F a través de S por

$$\Phi = \int_S F \cdot dS.$$

Notamos que el signo Φ cambia si reparametrizamos S, de modo que el vector normal determinado por la parametrización tiene la dirección opuesta.

Ejemplo 3. Supongamos que la velocidad de un flujo en $\mathcal{R}^3$ está dada por

$$F(x_1, x_2, x_3) = \left(\frac{x_1}{x_1^2 + x_2^2 + x_3^2}, \frac{x_2}{x_1^2 + x_2^2 + x_3^2}, \frac{x_3}{x_1^2 + x_2^2 + x_3^2}\right).$$

El flujo de F a través de la esfera S_a de centro en el origen y radio a es

$$\int_{S_a} F \cdot d\omega = \int_{S_a} \frac{x_1\, dx_2\, dx_3 + x_2\, dx_3\, dx_1 + x_3\, dx_1\, dx_2}{x_1^2 + x_2^2 + x_3^2}.$$

Sin embargo, el denominador es igual a a^2 en S_a, de modo que la integral de superficie toma la forma más simple

$$\frac{1}{a^2}\int_{S_a} x_1\, dx_2\, dx_3 + x_2\, dx_3\, dx_1 + x_3\, dx_1\, dx_2.$$

Podemos representar S_a paramétricamente en la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x^3 \end{pmatrix} = \begin{pmatrix} a \operatorname{sen} \varphi \cos \theta \\ a \operatorname{sen} \varphi \operatorname{sen} \theta \\ a \cos \varphi \end{pmatrix}, \qquad \begin{aligned} 0 \le \varphi \le \pi \\ 0 \le \theta \le 2\pi, \end{aligned}$$

de modo que

$$\frac{\partial(x_2, x_3)}{\partial(\varphi, \theta)} = a^2 \operatorname{sen}^2 \varphi \cos \theta$$

$$\frac{\partial(x_3, x_1)}{\partial(\varphi, \theta)} = a^2 \operatorname{sen}^2 \varphi \operatorname{sen} \theta$$

$$\frac{\partial(x_1, x_2)}{\partial(\varphi, \theta)} = a^2 \operatorname{sen} \varphi \cos \varphi.$$

Estas son las coordenadas de un vector normal que apunta hacia afuera de la esfera. Entonces

$$\int_{S_a} F \cdot dS = a\int_0^{2\pi} d\theta \int_0^{\pi} \operatorname{sen}^3 \varphi \cos^2 \theta + \operatorname{sen}^3 \varphi \operatorname{sen}^2 \theta + \operatorname{sen} \varphi \cos^2 \varphi\, d\varphi$$

$$= a\int_0^{2\pi} d\theta \int_0^{\pi} \operatorname{sen} \varphi\, d\varphi = 4\pi a.$$

Para calcular la integral de línea a lo largo de una curva lisa por pedazos es natural orientar los pedazos lisos, de modo que en el punto final de uno de ellos sea igual al punto final del siguiente. Para integrar un campo vectorial sobre una superficie lisa por pedazos necesitamos una noción de orientación para pedazos de superficie lisa. Si $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$ representa paramétricamente un pe-

dazo de superficie lisa, donde g está definida en D, entonces la Fig. 13 ilustra la relación entre D y S. El borde de S, que corresponde bajo g a la frontera de D, será llamado la **frontera** de S. A medida que un punto $\mathbf{u}$ se mueve a lo largo de la frontera lisa por pedazos de D en el sentido *opuesto* al de los punteros del reloj, su imagen $g(\mathbf{u})$ recorre la frontera de S, con lo que llamaremos **orientación positiva**. Otra interpretación geométrica de la orientación positiva se obtiene observando que cuando el vector normal $\partial g/\partial u \times \partial g/\partial v$ se mueve a lo largo de la frontera en la dirección positiva, la superficie S se mantiene a la izquierda.

Una **superficie lisa por pedazos** es una unión de pedazos de superficie lisa unidos a lo largo de las curvas frontera comunes. Una superficie lisa por pedazos es **orientable** si, al orientar positivamente las curvas fronteras de sus pedazos, las fronteras comunes a dos pedazos tienen orientaciones opuestas. Si ninguna de las dos curvas se invierte, sus parametrizaciones deben ser equivalentes. En la Fig. 18 se supone que las superficies son representables como superficies lisas por pedazos y que las dos primeras son orientables. Sin embargo, al unir estos dos pedazos, torciendo uno de ellos, se obtiene una cinta de Möbius que es un ejemplo típico de una superficie no orientable.

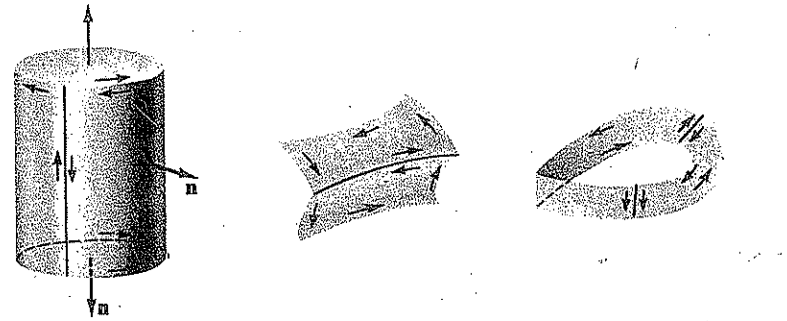


Figura 18

Definimos la integral de un campo vectorial continuo sobre una superficie lisa por pedazos como la suma de las integrales sobre cada uno de sus pedazos lisos. Por tanto, si $S = S_1 \cup S_2$,

$$\int_S F \cdot dS = \int_{S_1} F \cdot dS + \int_{S_2} F \cdot dS.$$

Esta definición es válida aun en el caso en que S no es orientable, pero, en la práctica, es de poco interés integrar campos vectoriales sobre superficies no orientables. Por otra parte, la integral de una función real sobre una superficie puede ser calculada sin referencia a su orientación. La razón para esto es que, en las Ecs. 3.1 y 3.2, el así llamado elemento de área,

$$d\sigma = \left|\frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v}\right| du\, dv,$$

no cambia cuando se invierte la orientación si intercambiamos u y v. Pero en la Ec. 3.3 el vector elemento de superficie

$$dS = \left(\frac{\partial g}{\partial u} \times \frac{\partial g}{dv}\right) du\, dv,$$

cambia de signo cuando se intercambia u y v. Observamos que el elemento de superficie dS también puede ser expresado en la forma

$$dS = \mathbf{n}\, d\sigma,$$

donde $\mathbf{n}$ es el vector unitario normal a la superficie dado por

$$\mathbf{n} = \frac{\dfrac{\partial g}{\partial u} \times \dfrac{\partial g}{\partial v}}{\left|\dfrac{\partial g}{\partial u} \times \dfrac{\partial g}{\partial v}\right|}.$$

EJERCICIOS

1. (a) Encuentre el área de la rampa espiral representada por

$$g(u, v) = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \\ v \end{pmatrix}, \qquad 0 \le u \le 1, \qquad 0 \le v \le 3\pi.$$

$$\left[\textit{Resp}: \frac{3\pi}{2}(\sqrt{2} + \log(1 + \sqrt{2})).\right]$$

(b) Suponga que la superficie de la parte (a) tiene una densidad por unidad de área en cada punto igual a la distancia del punto al eje central de la superficie. Encuentre la masa total de la superficie.

2. Calcule $\int_S F \cdot dS$, donde

(a) $F(x, y, z) = (x, y, z)$ y S está dada por

$$g(u, v) = \begin{pmatrix} u - v \\ u + v \\ uv \end{pmatrix}, \qquad \begin{matrix} 0 \le u \le 1, \\ 0 \le v \le 2. \end{matrix}$$

(b) $F(x, y, z) = (x^2, 0, 0)$ y S está dada por

$$g(u, v) = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \\ v \end{pmatrix}, \qquad \begin{matrix} 0 \le u \le 1, \\ 0 \le v \le 2\pi. \end{matrix}$$

3. Encuentre la masa total de una película esférica cuya densidad en cada punto es igual a la distancia del punto a un punto fijo de la esfera.

4. Sean $\mathbf{x} = g(u, v)$, para (u, v) en D y $\mathbf{x} = h(s, t)$ en B dos parametrizaciones para el mismo pedazo de superficie lisa S en $\mathcal{R}^3$. Si existe una transformación inyectiva T de D sobre B tal que T y T^{-1} son continuamente diferenciables con determinante jacobiano positivo y tal que $g(u, v) = h(T(u, v))$ para (u, v) en D, entonces g y h se llaman parametrizaciones **equivalentes** de S.

(a) Demuestre que dos parametrizaciones equivalentes asocian la misma área a S. [*Indicación*: Use el teorema del cambio de variables.]

(b) Demuestre que dos parametrizaciones equivalentes asocian la misma integral de superficie a un campo vectorial en S.

5. Suponga que la temperatura en un punto (x, y, z) de una región R está dada por una función continuamente diferenciable $T(x, y, z)$. Entonces el campo vectorial ∇T se llama el **gradiente de temperatura** y, bajo ciertas hipótesis físicas, $\nabla T(x, y, z)$ es proporcional a la dirección y flujo del calor por unidad de área en (x, y, z).

(a) Si $T(x, y, z) = x^2 + y^2$ para $x^2 + y^2 \le 4$, encuentre el flujo total de calor a través de la superficie cilíndrica $x^2 + y^2 = 1$, $0 \le z \le 1$.

(b) Dé un ejemplo de un campo vectorial continuamente diferenciable que no puede ser un gradiente de temperatura.

6. La función potencial newtoniano $(x^2 + y^2 + z^2)^{-1/2}$ tiene como gradiente el campo de fuerza atractiva F de una partícula cargada en el origen que actúa sobre una partícula de carga opuesta ubicada en (x, y, z). El flujo del campo a través de un pedazo de superficie lisa es, por definición, la integral de superficie de F sobre S. Demuestre que el flujo de F a través de una esfera de centro en el origen y radio a es independiente de a.

7. (a) Si $\mathcal{R}^2 \xrightarrow{f} \mathcal{R}$ es continuamente diferenciable en un conjunto D acotado por un pedazo de superficie lisa, demuestre que el área del gráfico de f es

$$\sigma(S) = \int_D \sqrt{1 + (f_x)^2 + (f_y)^2}\, dx\, dy.$$

(b) Encuentre el área del gráfico de $f(x, y) = x^2 + y$ para $0 \le x \le 1$, $0 \le y \le 1$.
[*Resp*. $\sqrt{\frac{3}{2}} + \frac{1}{2}\log(\sqrt{2} + \sqrt{3})$.]

8. Demuestre que si $\mathcal{R}^3 \xrightarrow{G} \mathcal{R}$ es continuamente diferenciable y determina implícitamente un pedazo de superficie lisa S en la cual $\partial G/\partial z \neq 0$ y que yace sobre una región D del plano xy, entonces

$$\sigma(S) = \int_D \sqrt{\left(\frac{\partial G}{\partial x}\right)^2 + \left(\frac{\partial G}{\partial y}\right)^2 + \left(\frac{\partial G}{\partial z}\right)^2} \left|\frac{\partial G}{\partial z}\right|^{-1} dx\, dy.$$

Suponga que hay un solo punto de S sobre cada punto de D.

9. Demuestre que la frontera de un pedazo de superficie lisa es una curva lisa por pedazos.

10. Encuentre una parametrización como superficie lisa por pedazos orientable con normal apuntando hacia afuera para cada uno de los conjuntos siguientes:

(a) El cilindro dado por $x^2 + y^2 = 1$, $0 \le z \le 1$ y $x^2 + y^2 \le 1$, $z = 0$.

(b) El embudo dado por $x^2 + y^2 - z^2 = 0$, $1 \le z \le 4$ y $x^2 + y^2 = 1$, $0 \le z \le 1$.

(c) La artesa dada por $y - z = 0$, $0 \le x \le 1$, $0 \le z \le 1$ e $y + z = 0$, $0 \le x \le 1$, $0 \le z \le 1$.

11. Sea F el campo vectorial en $\mathcal{R}^3$ dado por $F(x, y, z) = (x, y, 2z - x - y)$. Encuentre la integral de F sobre las superficies orientadas del Ejercicio 10.

12. Sea F un campo de velocidad de flujo continuo de un fluido y sea M una cinta de Möbius lisa por pedazos en el dominio de F. ¿Es posible definir el flujo de F a través de M?

13. Parametrice el conjunto del Ejercicio 10(a) de modo que no sea orientado y que sus normales apunten hacia afuera en la parte de abajo y en los costados. Calcule la integral de $F(x, y, z) = (x, y, 2z - x - y)$ sobre la superficie no orientada.

14. Demuestre que si F y G son campos continuos de vectores en una superficie lisa por pedazos S, entonces

$$\int_S (aF + bG) \cdot dS = a\int_S F \cdot dS + b\int_S G \cdot dS,$$

donde a y b son constantes.

15. (a) Sea F un campo vectorial continuo en un pedazo de superficie lisa S. Demuestre que

$$\left|\int_S F \cdot dS\right| \le M\sigma(S),$$

donde M es el máximo de $|F(\mathbf{x})|$ para $\mathbf{x}$ en S. [*Indicación*: Escriba $\int F \cdot dS$ en la forma $\int F \cdot \mathbf{n}\, d\sigma$.]

(b) Use la parte (a) para demostrar que si S se contrae a un punto $\mathbf{x}_0$ en forma tal que $\sigma(S)$ tiende a cero, entonces $(1/\sigma(S))\int_S F \cdot dS$ tiende a $F(\mathbf{x}_0) \cdot \mathbf{n}_0$, donde $\mathbf{n}_0$ es el vector unitario normal a S en $\mathbf{x}_0$.

16. Sea f una función real continuamente diferenciable de una variable, no negativa para $a \le x \le b$. El gráfico de f rotado alrededor del eje x genera una superficie de revolución en $\mathcal{R}^3$.

(a) Encuentre una parametrización de S en términos de f.

(b) Demuestre que $\sigma(S) = 2\pi \int_a^b f(x)\sqrt{1 + (f'(x))^2}\, dx$.

17. El **ángulo sólido** determinado por un cono sólido $\mathcal{C}$ en $\mathcal{R}^3$ con vértice en el origen es, por definición, el área de la intersección de $\mathcal{C}$ con la esfera unitaria $|\mathbf{x}| = 1$. Véase la Fig. 19.

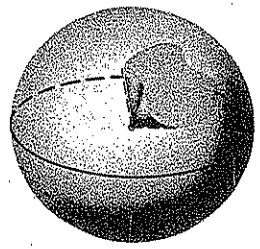


Figura 19

(a) Demuestre que una reducción adecuada de esta definición conduce a la definición habitual del ángulo entre dos rectas.

(b) Calcule el ángulo sólido determinado por el cono $x^2 + y^2 \le 2z^2$, $0 \le z$. [*Resp.* $2\pi(1 - 3^{-1/2})$.]

4. El teorema de Stokes

Se puede hacer una extensión importante del teorema de Green como sigue: en lugar de considerar una región plana D acotada por una curva, podemos pensar en elevar esta región, junto con su curva frontera a una superficie S en $\mathcal{R}^3$. Entonces S tendrá como frontera una curva γ en el espacio correspondiente a la frontera de D. Esta elevación se precisa definiendo en D y en su frontera lisa por pedazos una función $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$ tal que S sea la imagen de D. La Fig. 20

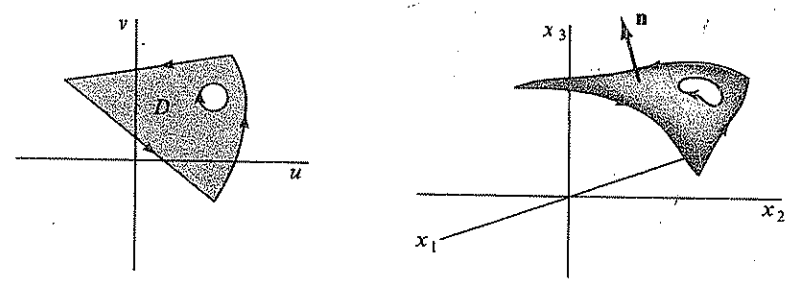


Figura 20

representa un caso típico. La región D tiene su frontera orientada en el sentido opuesto al de los punteros del reloj y la frontera γ de S queda así orientada positivamente con respecto de S. Si parametrizamos la frontera de D por $\mathcal{R} \xrightarrow{h} \mathcal{R}^2$, para $a \le t \le b$, entonces la función compuesta $g(h(t))$ describe la frontera de S. Denotaremos por ∂S la frontera de S orientada positivamente.

Ahora podemos expresar la relación entre la integral de línea de un campo vectorial F a lo largo de ∂S y la integral de superficie sobre S de un campo vectorial asociado a F. Suponemos que $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}^3$ es un campo vectorial continuamente diferenciable, cuyo dominio contiene a S. Entonces el campo vectorial **rot F** está definido por

$$\operatorname{rot} F(\mathbf{x}) = \left(\frac{\partial F_3}{\partial x_2}(\mathbf{x}) - \frac{\partial F_2}{\partial x_3}(\mathbf{x}), \frac{\partial F_1}{\partial x_3}(\mathbf{x}) - \frac{\partial F_3}{\partial x_1}(\mathbf{x}), \frac{\partial F_2}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_2}(\mathbf{x})\right),$$

donde F_1, F_2 y F_3 son las funciones coordenadas de F. Si el dominio de F es un conjunto abierto, entonces el dominio de rot F es el mismo conjunto. El campo vectorial rot F ha sido elegido de manera tal que si S es un pedazo suficientemente liso de superficie, entonces

$$\int_S \operatorname{rot} F \cdot d\omega = \oint_{\partial S} F \cdot d\mathbf{x}. \tag{1}$$

Nótese que si F fuera esencialmente un campo vectorial de dimensión 2, con F_3 idénticamente cero y F_1 y F_2 independientes de x_3, distinta de cero y la Ec. (1) se reduciría a la fórmula de Green.

Ejemplo 1. Sea S la superficie espiral parametrizada por

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} u\cos v \\ u \operatorname{sen} v \\ v \end{pmatrix}, \quad \text{para} \quad 0 \le u \le 1, \qquad 0 \le v \le \frac{\pi}{2}.$$

Entonces la frontera de S está formada por tres segmentos de recta y una curva espiral, tal como lo muestra la Fig. 21. Restringiendo la parametrización de

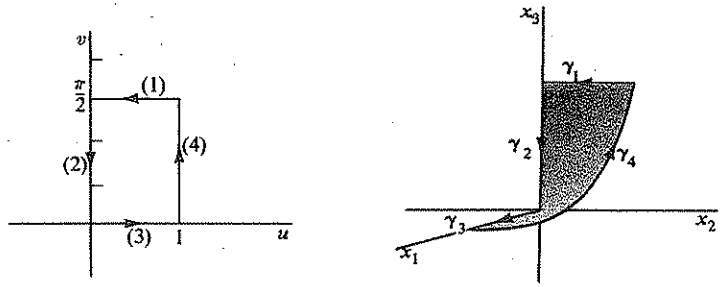


Figura 21

S a la frontera de D se obtienen las siguientes parametrizaciones para los pedazos lisos de la frontera de S.

$$\gamma_1: \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1-t \\ \pi/2 \end{pmatrix}, \quad 0 \le t \le 1$$

$$\gamma_2: \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ (\pi/2)-t \end{pmatrix}, \quad 0 \le t \le \frac{\pi}{2}$$

$$\gamma_3: \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix}, \quad 0 \le t \le 1$$

$$\gamma_4: \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos t \\ \operatorname{sen} t \\ t \end{pmatrix}, \quad 0 \le t \le \frac{\pi}{2}.$$

Ahora, sea F el campo vectorial $F(x_1, x_2, x_3) = (x_1, x_2, x_3)$. La integral de línea de F a lo largo de γ_i tiene la forma

$$\int_{\gamma_i} x_3\,dx_1 + x_1\,dx_2 + x_2\,dx_3.$$

Es fácil ver que las integrales a lo largo de γ_1, γ_2 y γ_3 son cero, mientras que a lo largo de γ_4 tenemos que

$$\int_{\gamma_4} F \cdot d\mathbf{x} = \int_0^{\pi/2} (\cos^2 t + \operatorname{sen} t - t \operatorname{sen} t)\,dt$$
$$= \frac{\pi}{4}.$$

Por otra parte, rot $F(x_1, x_2, x_3) = (1, 1, 1)$, de modo que la integral de rot F sobre S es

$$\int_S \operatorname{rot} F \cdot dS = \int_D \left(\frac{\partial(x_2, x_3)}{\partial(u, v)} + \frac{\partial(x_3, x_1)}{(u, v)} + \frac{\partial(x_1, x_2)}{\partial(u, v)} \right) du\,dv$$
$$= \int_0^1 du \int_0^{\pi/2} (\operatorname{sen} v - \cos v + u)\,dv = \frac{\pi}{4}.$$

Esto verifica la Ec. (1) en este caso especial.

La demostración que daremos del teorema de Stokes depende de una aplicación del teorema de Green a la región D, en la cual está definida la parametrización de S. Por esta razón es necesario hacer hipótesis adecuadas para D, de modo que sea válido el teorema de Green. Además, si $\mathfrak{R}^2 \xrightarrow{g} \mathfrak{R}^3$ es la parametrización de S, necesitamos que las segundas derivadas de g sean continuas en D. Estas condiciones pueden ser relajadas a costa de hacer mucho más difícil la demostración.

4.1 Teorema de Stokes

Sea S un pedazo de superficie lisa en $\mathfrak{R}^3$ parametrizado por una función g con segundas derivadas continuas. Supongamos que el dominio D de g es una unión finita de regiones simples acotadas por curvas lisas por pedazos. Si F es un campo vectorial diferenciable definido en S, entonces

$$\int_S \operatorname{rot} F \cdot dS = \oint_{\partial S} F \cdot d\mathbf{x},$$

donde ∂S es la frontera de S orientada positivamente.

Demostración. Sean F_1, F_2, F_3 las funciones coordenadas de F. Demostraremos que

$$\oint_{\partial S} F_1\,dx_1 = \int_S -\frac{\partial F_1}{\partial x_2}\,dx_1\,dx_2 + \frac{\partial F_1}{\partial x_3}\,dx_3\,dx_1. \tag{2}$$

Las demostraciones de que

$$\oint_{\partial S} F_2\,dx_2 = \int_S -\frac{\partial F_2}{\partial x_3}\,dx_2\,dx_3 + \frac{\partial F_2}{dx_1}\,dx_1\,dx_2$$

y

$$\oint_{\partial S} F_3\,dx_3 = \int_S -\frac{\partial F_3}{\partial x_1}\,dx_3\,dx_1 + \frac{\partial F_3}{dx_2}\,dx_2\,dx_3$$

son análogas y la suma de estas tres ecuaciones da el teorema de Stokes. Para demostrar la primera ecuación, supongamos que $h(t) = (u(t), v(t))$ es una parametrización orientada en el sentido opuesto al de los punteros del reloj de la frontera δ de D. Entonces $g(h(t))$ es una parametrización lisa por pedazos y orientada positivamente de la frontera de S. Si denotamos por g_1, g_2, g_3 las funciones coordenadas de g, tenemos que

$$\begin{aligned}\oint_{\partial S} F_1\,dx_1 &= \int F_1(g(u,v))\frac{d}{dt}\,g_1(u,v)\,dt\\ &= \int F_1(g(u,v))\left[\frac{\partial g_1}{\partial u}(u,v)\frac{du}{dt} + \frac{\partial g_1}{\partial v}(u,v)\frac{dv}{dt}\right]dt\\ &= \oint_{\delta} F_1 \circ g\,\frac{\partial g_1}{\partial u}\,du + F_1 \circ g\,\frac{\partial g_1}{\partial v}\,dv.\end{aligned}$$

La última integral es una integral de línea sobre la región D en $\mathcal{R}^2$ y podemos aplicar el teorema de Green, obteniendo

$$\oint_{\partial S} F_1\,dx_1 = \int_D \left[\frac{\partial}{\partial u}\left(F_1 \circ g\,\frac{\partial g_1}{\partial v}\right) - \frac{\partial}{\partial v}\left(F_1 \circ g\,\frac{\partial g_1}{\partial u}\right)\right] du\,dv. \tag{3}$$

Como g tiene segundas derivadas continuas, la integral sobre D existe. Esto mismo hace posible el intercambio del orden de diferenciación en las derivadas parciales, de modo que

$$\frac{\partial}{\partial u}\left(F_1 \circ g\,\frac{\partial g_1}{\partial v}\right) - \frac{\partial}{\partial v}\left(F_1 \circ g\,\frac{\partial g_1}{\partial u}\right) = -\frac{\partial F_1}{\partial x_2}\frac{\partial(g_1, g_2)}{\partial(u, v)} + \frac{\partial F_1}{\partial x_3}\frac{\partial(g_3, g_1)}{\partial(u, v)}. \tag{4}$$

La Ec. (2) se obtiene sustituyendo esta identidad en la Ec. (3), con lo cual se completa la demostración.

Mediante el teorema de Stokes podemos obtener una interpretación del campo vectorial rot F, que proporciona información acerca de F. Sea $\mathbf{x}_0$ un punto de un conjunto abierto, en el cual F es continuamente diferenciable. Sea $\mathbf{n}_0$ un vector unitario arbitrario que apunta hacia afuera en $\mathbf{x}_0$ y construyamos un disco S_r de radio r y centro en $\mathbf{x}_0$, perpendicular a $\mathbf{n}_0$. Esta construcción está indicada en la Fig. 22. Aplicando el teorema de Stokes a F en la superficie S_r y su frontera γ_r obtenemos

$$\oint_{\gamma_r} F \cdot d\mathbf{x} = \int_S \text{rot } F \cdot dS.$$

El valor de esta integral de línea se llama la **circulación** de F a lo largo de γ_r y mide la fuerza del campo en la dirección tangente a γ_r. Por tanto, si r es pequeño la circulación a lo largo de γ_r es una medida de la tendencia del campo cerca de $\mathbf{x}_0$, a rotar alrededor del eje determinado por $\mathbf{n}_0$. Por otra parte, la integral de superficie, para r suficientemente pequeño, es aproximadamente igual al producto punto rot $F(\mathbf{x}_0) \cdot \mathbf{n}_0$ multiplicado por el área de S_r. Véase el Ejercicio 15 de la sección anterior. Por tanto, la circulación a lo largo de γ_r será mayor si $\mathbf{n}_0$ tiene la misma dirección de rot $F(\mathbf{x}_0)$. Por tanto, podemos decir que rot $F(\mathbf{x}_0)$ determina el eje alrededor del cual la circulación es máxima cerca de $\mathbf{x}_0$. Análogamente, $|\text{rot } F(\mathbf{x}_0)|$ mide la magnitud de la circulación alrededor de este eje, cerca de $\mathbf{x}_0$.

La extensión del teorema de Stokes a superficies lisas por pedazos orientables es muy sencilla, a pesar de que hay que ser cuidadoso al definir la frontera de este tipo de superficie. La Fig. 23 ilustra el método adecuado. Las superficies

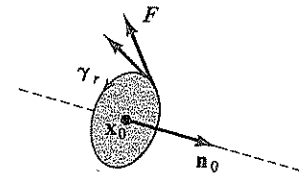


Figura 22

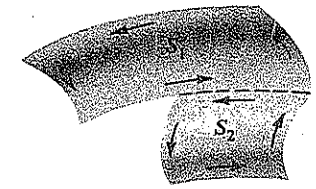


Figura 23

S_1 y S_2 tienen sus fronteras unidas de modo que ellas forman una superficie lisa por pedazos positivamente orientada que denotamos por $S_1 \cup S_2$. Recordemos que la integral de superficie de un campo vectorial F sobre $S_1 \cup S_2$ ha sido definido por

$$\int_{S_1 \cup S_2} F \cdot dS = \int_{S_1} F \cdot dS + \int_{S_2} F \cdot dS.$$

La parte común de frontera, representada por una línea punteada en la Fig. 23, será recorrida en direcciones opuestas si usamos las parametrizaciones inducidas por S_1 y S_2. Por tanto, las respectivas integrales de línea de F sobre la frontera común tendrán signos opuestos y cuando se sumen las integrales de línea a lo largo de ∂S_1 y ∂S_2, ellas se cancelarán dejando solo una integral de línea a lo largo del resto de las fronteras de S_1 y S_2. Esta parte restante se llama la frontera positivamente orientada de $S_1 \cup S_2$ y es denotada por $\partial(S_1 \cup S_2)$. (De modo que, en general, $\partial(S_1 \cup S_2) \neq \partial S_1 \cup \partial S_2$.) Con esta convención, escribimos el teorema de Stokes en la forma

$$\int_S \text{rot } F \cdot dS = \oint_{\partial S} F \cdot d\mathbf{x},$$

para una superficie lisa por pedazos S.

Ejemplo 2. Una esfera puede ser considerada como una superficie lisa por pedazos en la cual todas las curvas frontera se cancelan entre sí. En efecto, si parametrizamos una esfera S_a en $\mathcal{R}^3$ por

$$g(u, v) = \begin{pmatrix} a \operatorname{sen} v \cos u \\ a \operatorname{sen} v \operatorname{sen} u \\ a \cos v \end{pmatrix}, \qquad \begin{matrix} 0 \le u \le 2\pi \\ 0 \le v \le \pi, \end{matrix}$$

entonces la «frontera» positivamente orientada de la esfera está formada por el semicírculo mostrado en la Fig. 24 recorrida una vez en cada dirección. El semicírculo corresponde a los segmentos $u = 0$ y $u = 2\pi$ en el dominio. (¿Qué sucede con los segmentos $v = 0$ y $v = \pi$?) El resultado es que la integral de línea a lo largo de ∂S_a es cero y del teorema de Stokes, aplicado al campo vectorial F en S_a, se tiene que

$$\int_{S_a} \operatorname{rot} F \cdot dS = 0.$$

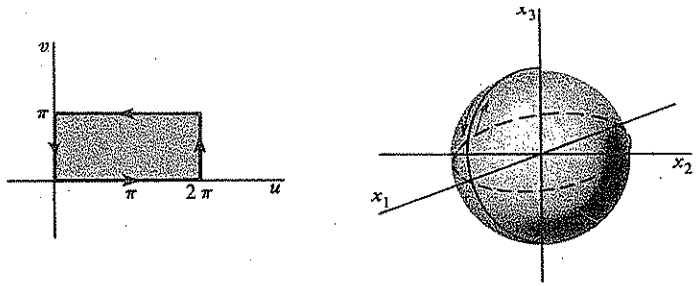


Figura 24

Una superficie S tal como la esfera en la cual la frontera no tiene ningún efecto en el cálculo de la integral de línea a lo largo de S se llama una **superficie cerrada**.

El teorema de Stokes da información interesante acerca de los campos gradientes; es decir, campos F tales que $F = \nabla f$ o, usando otra notación, $F = \operatorname{grad} f$. Si suponemos que F es el campo continuamente diferenciable gradiente de f, podemos formar el campo vectorial rot F. Como $F = (\partial f/\partial x_1, \partial f/\partial x_2, \partial f/\partial x_3)$, de la definición de rot F y de la igualdad de las derivadas parciales mixtas, se tiene que

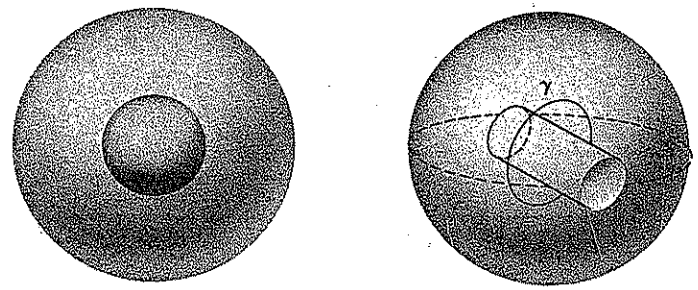
$$\text{4.2} \qquad \operatorname{rot} (\operatorname{grad} f)(\mathbf{x}) = 0,$$

para todo $\mathbf{x}$ en el dominio de f. En los Teoremas 2.5 y 2.6 hemos encontrado la condición rot $F = 0$ expresada en términos de la matriz jacobiana

$$F' = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_2}{\partial x_1} & \dfrac{\partial F_3}{\partial x_1} \\ \dfrac{\partial F_1}{\partial x_2} & \dfrac{\partial F_2}{\partial x_2} & \dfrac{\partial F_3}{\partial x_2} \\ \dfrac{\partial F_1}{\partial x_3} & \dfrac{\partial F_2}{\partial x_3} & \dfrac{\partial F_3}{\partial x_3} \end{pmatrix}.$$

Es claro que la simetría de F' con respecto de la diagonal principal es equivalente a rot $F = 0$. En particular, el Teorema 2.5 dice que si F es un campo gradiente, entonces rot F es idénticamente cero. El Teorema 2.6 es un converso parcial y dice que si rot F es idénticamente cero, entonces existe un rectángulo en el cual F es un campo gradiente. A veces esto se expresa diciendo que F es *localmente* un campo gradiente. El Ejemplo 2 de la Sec. 2 muestra que el converso estricto es falso en general. Mediante el teorema de Stokes podemos demostrar otro converso parcial, en el cual la condición local se remplaza por un tipo diferente de restricción en el dominio del campo dado.

Para esto definimos lo que se entiende por un conjunto abierto **simplemente conexo** B en $\mathcal{R}^n$. A grandes trazos, un conjunto B es simplemente conexo si toda curva cerrada γ en B puede ser contraída a un punto manteniéndose siempre dentro de B. A medida que γ se contrae a un punto, se barre una superficie S en B cuya frontera es γ. La región entre las dos esferas mostradas en la Fig. 25(a)



(a) (b)

Figura 25

es simplemente conexa. Sin embargo, la bola abierta con un agujero que la atraviesa no es simplemente conexa, ya que toda superficie cuya frontera circunda el agujero debe estar por lo menos parcialmente fuera de B. Una región simplemente conexa en $\mathcal{R}^2$ es el interior de una curva cerrada, mientras que su exterior no lo es. En la Fig. 26(a), la curva γ es la frontera de la superficie formada por la parte del plano en el interior de γ. Sin embargo, la presencia de un agujero en la Fig. 26(b) impide una construcción similar. En forma más precisa, diremos que un conjunto abierto es simplemente conexo si toda curva cerrada lisa por pedazos γ en B es la frontera de alguna superficie lisa por pe-

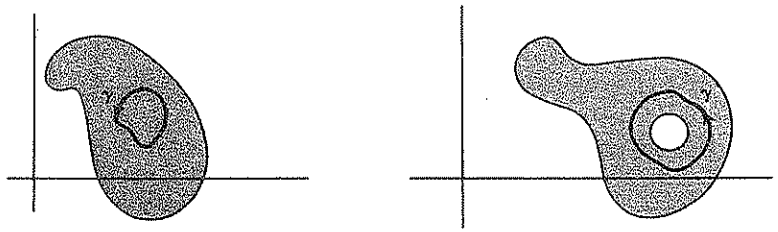


Figura 26

dazos orientable S cuyo dominio de parametrización es un disco en $\mathcal{R}^2$. En las aplicaciones, suponemos que S está parametrizada por una función con segundas derivadas parciales continuas. Generalmente, no se incluyen las condiciones de diferenciabilidad en la definición de conjunto simplemente conexo, exigiéndose solo la continuidad. Sin embargo, como no usaremos ese concepto más general, no le damos un nombre especial.

Ahora podemos demostrar el siguiente

4.3 Teorema

Sea F un campo continuamente diferenciable de vectores definido en un conjunto abierto B en $\mathcal{R}^2$ o $\mathcal{R}^3$. Si

1. B es simplemente conexo, y
2. rot F es idénticamente cero en B,

entonces F es un campo gradiente en B, es decir, existe una función real f tal que $F = \nabla f$.

Demostración. Por el Teorema 2.4, es suficiente demostrar que $\oint_\gamma F \cdot d\mathbf{x} = 0$ para toda curva lisa por pedazos γ en B. Como B es simplemente conexo, existe una superficie lisa por pedazos S cuya frontera es γ y a la cual podemos aplicar el teorema de Stokes en 2 o 3 dimensiones. Por tanto,

$$\oint_{\partial S} F \cdot d\mathbf{x} = \int_S \text{rot } F \cdot dS = 0.$$

EJERCICIOS

1. Calcule rot F si
   (a) $F(x_1, x_2, x_3) = (x_2 - x_3^2, x_3 - x_1^2, x_1 - x_2^2)$.
   (b) $F(x, y, z) = (x, 2y, 3z)$.
2. Verifique el teorema de Stokes calculando las dos integrales para el campo vectorial $F(x_1, x_2, x_3) = (x_1, x_2, x_3)$ en un hemisferio de centro en el origen en $\mathcal{R}^3$.

3. (a) Verifique que si $F(x_1, x_2, x_3)$ es independiente de x_3 y la tercera coordenada de F es idénticamente cero entonces la fórmula de Stokes, aplicada a una superficie en el plano x_1, x_2, es igual a la fórmula de Green.
   (b) Considere la función $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$ definida por

$$g(u, v) = \begin{pmatrix} u \cos v \\ u \operatorname{sen} v \\ 0 \end{pmatrix}, \qquad \begin{matrix} 1 \le u \le 2, \\ 0 \le v \le 4\pi. \end{matrix}$$

   Si S es la imagen de g, dé una descripción precisa de la frontera de S.
   (c) Si F es un campo vectorial continuamente diferenciable en S del tipo descrito en la parte (a), use el teorema de Stokes para calcular la integral de F sobre la frontera positivamente orientada de S.
4. Demuestre que la fórmula de Stokes puede ser escrita en la forma

$$\int_S \text{rot } F \cdot \mathbf{n}\, d\sigma = \oint_{\partial S} F \cdot \mathbf{t}\, ds,$$

   donde $\mathbf{n}$ es un vector unitario normal a S y $\mathbf{t}$ un vector unitario tangente a ∂S.
5. Use el resultado del Ejercicio 15 de la sección anterior y el teorema de Stokes para demostrar que si F es un campo vectorial continuamente diferenciable y $\mathbf{x}_0$ es un punto en el dominio de definición de F, entonces

$$\lim_{r \to 0} \frac{1}{A(D_r)} \oint_c F \cdot \mathbf{t}\, ds = \text{rot } F(\mathbf{x}_0) \cdot \mathbf{n}_0,$$

   donde D_r es un disco de radio r de centro $\mathbf{x}_0$ y $\mathbf{n}_0$ es un vector unitario normal al disco y c es la frontera de D.
6. Demuestre que si F es un campo vectorial continuamente diferenciable tal que en cada punto $\mathbf{x}$ de un pedazo de superficie lisa S el vector rot $F(\mathbf{x})$ es tangente a S, entonces la integral de F a lo largo de la frontera de S es cero.
7. Sea F un campo vectorial diferenciable definido en un subconjunto abierto B de $\mathcal{R}^3$. Use la descomposición de una matriz A en una matriz simétrica y una antisimétrica dada por $A = \frac{1}{2}(A + A') + \frac{1}{2}(A - A')$, para demostrar que para todo $\mathbf{y}$ en $\mathcal{R}^3$,

$$F'(\mathbf{x})\mathbf{y} = S(\mathbf{x})\mathbf{y} + \tfrac{1}{2} \text{ rot } F(\mathbf{x}) \times \mathbf{y},$$

   donde $S(\mathbf{x})$ es una matriz simétrica.
8. Sea F el campo gradiente del potencial newtoniano $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$. Demuestre que la circulación de F es cero en la vecindad de todo punto del dominio de F.
9. Efectúe los cálculos necesarios para demostrar la identidad de la Ec. (4) de la demostración del teorema de Stokes.
10. Considere el bote cilíndrico C de radio 1 con una frontera lisa no especificada y una orientación tal como la de la Fig. 27. Sea $F(x_1, x_2, x_3) = 2x_2^2, x_1^2, 3x_3^2)$. ¿Cuál es el valor de la integral de línea de F a lo largo de la frontera de C?
11. Calcule la integral de rot F, donde $F(x_1, x_2, x_3) = (x_2^3, -x_1^3, x_3^3)$ sobre el hemisferio

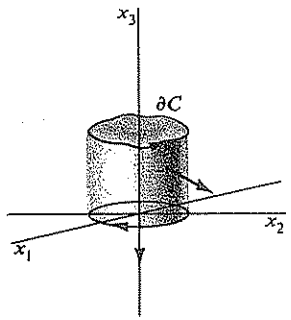


Figura 27

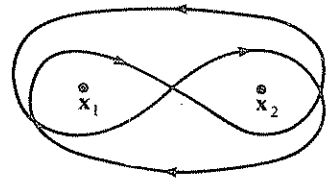


Figura 28

$x_1^2 + x_2^2 + x_3^2 = 1$, $x_3 \geq 0$, considerando una integral sobre el disco que cierra el hemisferio.

12. Demuestre que el subconjunto abierto de $\mathfrak{R}^2$ formado por todos los puntos distintos del origen no es simplemente conexo. Para ello encuentre un campo vectorial F para el cual rot F es idénticamente cero, pero tal que F no es un campo gradiente. [*Indicación*: Véase el Ejercicio 7 de la Sec. 1.]

13. El conjunto abierto en $\mathfrak{R}^2$ formado por $\mathfrak{R}^2$ con los dos puntos $\mathbf{x}_1$ y $\mathbf{x}_2$ removidos, no es simplemente conexo. Sin embargo, demuestre que si F es un campo vectorial continuamente diferenciable tal que rot $F = 0$ en la región mencionada, entonces la integral de F a lo largo de la curva lisa mostrada en la Fig. 28 es cero.

14. Sea B el conjunto abierto en R^3 formado por todos los puntos distintos del origen. Demuestre que un círculo con centro en el origen es la frontera de una superficie lisa por pedazos en B.

5. El teorema de Gauss

El teorema de Gauss tiene muchas aplicaciones, algunas de las cuales serán discutidas en la Sec. 6. Igual que el teorema de Stokes, puede ser considerado como una extensión del teorema de Green. Consideramos una región R en $\mathfrak{R}^3$ cuya frontera es una superficie lisa por pedazos S. Cada pedazo de S es parametrizado por una función diferenciable $\mathfrak{R}^2 \xrightarrow{g} \mathfrak{R}^3$ tal que el vector normal $\partial g/\partial u \times \partial g/\partial v$ está dirigido hacia afuera de R en cada punto de S. Entonces se dice que la superficie frontera tiene **orientación positiva** y, con esta orientación, la denotaremos ∂R. Para enunciar el teorema consideramos un campo vectorial F, continuamente diferenciable en R y en su frontera. Definimos la **divergencia** de F como la función real div F definida en R por

$$\operatorname{div} F(\mathbf{x}) = \frac{\partial F_1}{\partial x_1}(\mathbf{x}) + \frac{\partial F_2}{\partial x_2}(\mathbf{x}) + \frac{\partial F_3}{\partial x_3}(\mathbf{x}),$$

donde F_1, F_2, F_3 son las funciones coordenadas de F. Entonces la fórmula de Gauss (o de la divergencia) es

$$\int_R \operatorname{div} F\, dV = \int_{\partial R} F \cdot dS.$$

La fórmula de Gauss es parecida a la fórmula de Stokes y de Green y a la fórmula

$$\int_a^b \operatorname{grad} f \cdot d\mathbf{x} = f(\mathbf{b}) - f(\mathbf{a}),$$

ya que relacionan la integral de algún tipo de derivada de una función con el comportamiento de la función en la frontera. En cada caso, la orientación de la frontera es importante. Por ejemplo, si aplicamos el teorema de Gauss a la región R en $\mathfrak{R}^3$ definida por $1 \leq |x| \leq 2$, entonces la frontera orientada, ∂R, debe ser tal que los vectores normales a la esfera exterior tienen la dirección que apunta del origen hacia afuera y los vectores normales a la esfera interior apuntan hacia el origen. Esta situación está ilustrada en la Fig. 29. Diremos que ∂R tiene **orientación positiva** con respecto de R si los vectores normales dados por la parametrización de ∂R apuntan hacia afuera de R.

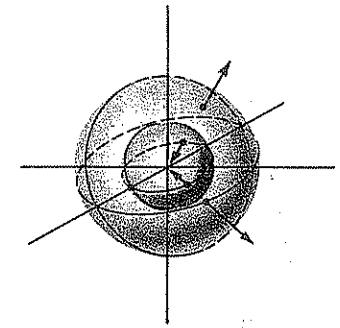


Figura 29

Demostraremos el teorema de Gauss para el caso en que R es una unión finita de regiones simples, donde se entiende por **región simple** en $\mathfrak{R}^3$ una región que es atravesada a lo más dos veces por toda recta paralela a un eje de coordenadas. La región comprendida entre dos esferas, mostrada en la Fig. 29, es una unión de ocho regiones simples, una en cada octante.

5.1 Teorema de Gauss

Sea R una unión finita de regiones simples en $\mathfrak{R}^3$ con una orientación positiva en su frontera lisa por pedazos ∂R. Si F es un campo vectorial diferenciable en R y ∂R, entonces

$$\int_R \operatorname{div} F\, dV = \int_{\partial R} F \cdot dS.$$

Demostración. En términos de las funciones coordenadas de F, la fórmula de Gauss tiene la forma

$$\int_R \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}\right) dx_1\, dx_2\, dx_3 = \int_{\partial R} F_1\, dx_2\, dx_3 + F_2\, dx_3\, dx_1 + F_3\, dx_\alpha\, dx_2.$$

Primero suponemos que R es una región simple y demostramos la ecuación

$$\int_R \frac{\partial F_2}{\partial x_2}\, dx_1\, dx_2\, dx_3 = \int_{\partial R} F_2\, dx_3\, dx_1,$$

ya que las demostraciones de la igualdad para los términos en F_1 y F_3 es similar. Sumando las ecuaciones resultantes, demuestre el teorema para regiones simples. Como R es simple, ∂R está formado por el gráfico de dos funciones $s(x_1, x_3)$ y $r(x_1, x_3)$ y, posiblemente, de pedazos paralelos al eje x_2, tal como lo muestra la Fig. 30. Sea

$$g(u, v) = \begin{pmatrix} g_1(u, v) \\ g_2(u, v) \\ g_3(u, v) \end{pmatrix}, \qquad (u, v) \text{ en } D$$

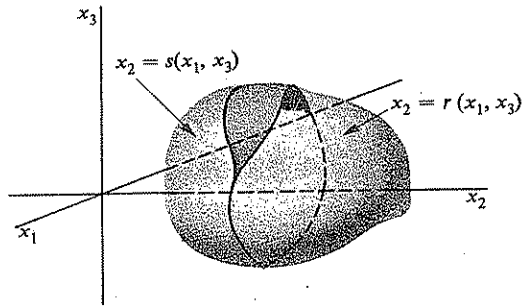


Figura 30

una parametrización de ∂R que la orienta positivamente. Entonces, por definición de la integral de superficie

$$\int_{\partial R} F_2\, dx_3\, dx_1 = \int_D F_2(g_1, g_2, g_3) \frac{\partial(g_3, g_1)}{\partial(u, v)}\, du\, dv, \tag{1}$$

y, en las partes de ∂R que son paralelas al eje x_2, el vector normal a ∂R es perpendicular al eje x_2. Por tanto, $\partial(g_3, g_1)/\partial(u, v)$, es decir, la tercera coordenada de la normal, es igual a cero, por lo cual se elimina la parte de la integral

que no es el gráfico de r o de s. Aplicamos ahora el teorema del cambio de variable a las dos partes restantes de la integral de la Ec. (1). La transformación apropiada es

$$\begin{pmatrix} x_3 \\ x_1 \end{pmatrix} = \begin{pmatrix} g_3(u, v) \\ g_1(u, v) \end{pmatrix},$$

con (u, v) en D_r o D_s, donde D_r y D_s son las partes de D que corresponden a los gráficos de r y de s. El determinante jacobiano $\partial(g_3, g_1)/\partial(u, v)$ es positivo en el gráfico de r y negativo en el gráfico de s, ya que él representa la coordenada x_2 de la normal que apunta hacia afuera. En D_r tenemos que $g_2(u, v) = r(x_1, x_3)$, mientras que en D_s, $g_2(u, v) = s(x_1, x_3)$. Usando estos hechos, del teorema del cambio de variable y de la Ec. (1), se tiene que

$$\int_{\partial R} F_2\, dx_3\, dx_\alpha = \int_{R_2} F_2(x_1, s(x_1, x_3), x_3)(-1)\, dx_1\, dx_3 + \int_{R_2} F_2(x_1, r(x_1, x_3), x_3)\, dx_1\, dx_3,$$

donde R_2 es la región plana que se obtiene proyectando R en el plano x_1x_3. Estas últimas integrales no son integrales de superficie, sino integrales sobre un conjunto. Entonces, por el teorema fundamental del cálculo,

$$\int_{\partial R} F_2\, dx_3\, dx_1 = \int_{R_2} \left[\int_{s(x_1,x_3)}^{r(x_1,x_3)} \frac{\partial F_2}{\partial x_2}(x_1, x_2, x_3)\, dx_2\right] dx_1\, dx_3 = \int_R \frac{\partial F_2}{\partial x_2}\, dx_1\, dx_2\, dx_3.$$

Argumentos análogos usando F_1 y F_3 completan la demostración para regiones simples, ya que sumando las tres ecuaciones así obtenidas, se tiene que

$$\int_{\partial R} F_1\, dx_2\, dx_3 + F_2\, dx_3\, dx_1 + F_3\, dx_1\, dx_2 = \int_R \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}\right) dx_1\, dx_2\, dx_3.$$

y esta es la fórmula de Gauss en coordenadas.

La extensión del teorema de Gauss a una unión finita R de regiones es esencialmente igual a la extensión análoga del teorema de Green. En el caso que nos ocupa, cuando dos regiones simples tienen una superficie de frontera en común, los vectores correspondientes serán negativos uno del otro. Las integrales de superficie correspondientes serán negativas una de otra y se cancelarán mutuamente. Las integrales de superficie que restan se extienden sobre la superficie ∂R.

Ejemplo 1. El Ejercicio 6 de la Sec. 3 consistió en mostrar que el flujo del campo gradiente F de la función potencial $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$

a lo largo de una esfera de radio a y centro en el origen, es independiente de a. Usando el teorema de Gauss, podemos demostrar algo más general y con un mínimo de cálculos. Sean S_1 y S_2 dos pedazos de superficie lisa cerradas cualesquiera, una contenida en la otra, tal que ambas contienen al origen y que forman la frontera de la región contenida entre ellas. Por ejemplo, R puede ser la región entre dos esferas. Un cálculo rutinario del gradiente muestra que $F(x, y, z) = (x^2 + y^2 + z^2)^{-3/2}(-x, -y, -z)$ y entonces la divergencia de este campo es cero, es decir, $\operatorname{div} F = 0$ en todas partes, salvo el origen. En particular, $\operatorname{div} F = 0$ en R. Aplicando el teorema de Gauss a R se obtiene

$$\int_{\partial R} F \cdot dS = \int_R \operatorname{div} F \, dV = 0.$$

Pero ∂R está formado por S_1 con la normal dirigida hacia adentro y S_2 con la normal dirigida hacia afuera, de modo que si S_1^- es la superficie interior con la normal opuesta, entonces

$$\int_{\partial R} F \cdot dS = -\int_{S_1^-} F \cdot dS + \int_{S_2} F \cdot dS = 0.$$

Por tanto, las integrales sobre las superficies orientadas hacia afuera, son iguales. Para encontrar el valor de estas integrales es suficiente calcularlas para una superficie, digamos una esfera. El resultado es -4π.

La divergencia de un campo vectorial F en un punto $\mathbf{x}$ puede ser interpretada como la tendencia del campo a radiar hacia afuera desde $\mathbf{x}$, lo cual es la razón para usar el término «divergencia». Para ver esto, consideremos la bola B_r de radio r y centro en un punto interior $\mathbf{x}_0$ del conjunto en el cual F es continuamente diferenciable. Dividiendo por $V(B_r)$ ambos términos de la fórmula de Gauss, se tiene que

$$\frac{1}{V(B_r)} \int_{B_r} \operatorname{div} F \, dV = \frac{1}{V(B_r)} \int_{\partial B} F \cdot \mathbf{n} \, d\sigma,$$

donde $\mathbf{n}$ es el vector unitario normal dirigido hacia afuera. Cuando r tiende a cero, el término de la izquierda tiende a $\operatorname{div} F(\mathbf{x}_0)$. Véase el Ejercicio 17 de la Sec. 2 del Cap. 6. En el término de la derecha se tiene la velocidad media de flujo por unidad de volumen a través de la esfera de radio r y centro en $\mathbf{x}_0$. Por tanto, el límite cuando r tiende a cero es, por definición, la velocidad de flujo hacia afuera por unidad de volumen en $\mathbf{x}_0$. Por esto, el teorema de Gauss se conoce también como el teorema de la divergencia.

EJERCICIOS

1. Calcule la divergencia de los campos vectoriales siguientes:
   (a) $F(x_1, x_2, x_3) = (x_1^2, x_2^2, x_3^2)$.
   (b) $F(x, y, z) = (\operatorname{sen} xy, 0, 0)$.
   (c) $F(x_1, x_2, x_3) = (x_2, x_3, x_1)$.

2. Demuestre las identidades siguientes para cualquier campo vectorial F o función real f dos veces continuamente diferenciable.
   (a) $\operatorname{div}(\operatorname{rot} F)(\mathbf{x}) = 0$.
   (b) $\operatorname{rot}(\operatorname{grad} f)(\mathbf{x}) = 0$.

3. (a) Demuestre que si $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$, entonces
   $$\operatorname{div}(\operatorname{grad} f)(\mathbf{x}) = 0 \text{ para todo } \mathbf{x} \neq 0.$$
   (b) Demuestre mediante un ejemplo que es posible que
   $$\operatorname{div}(\operatorname{grad} f)(\mathbf{x}) \neq 0$$
   para alguna función f dos veces continuamente diferenciable.
   (c) Si el operador Δ está definido por
   $$\Delta f = \operatorname{div}(\operatorname{grad} f),$$
   encuentre una fórmula para Δf en términos de las derivadas parciales de f. (Una función f tal que $\Delta f(\mathbf{x}) = 0$ para todo $\mathbf{x}$ en el dominio de f, se llama y Δ se llama el operador de Laplace.)

4. Por definición, la traza de una matriz cuadrada es la suma de los elementos en su diagonal principal. Si $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^n$ es un campo vectorial diferenciable, definimos la función real $\operatorname{div} F$ por
   $$\operatorname{div} F(\mathbf{x}) = \operatorname{tr} F'(\mathbf{x}),$$
   donde $\operatorname{tr} A$ denota la traza de A. Demuestre que en dimensiones 2 y 3, esta definición coincide con la que diéramos anteriormente.

5. Use el teorema de Gauss para calcular
   $$\int_S F \cdot dS$$
   sobre la esfera de radio 1 y centro en el origen de $\mathbb{R}^3$ con normal dirigida hacia afuera, si F es
   (a) $F(x_1, x_2, x_3) = (x_1^2, x_2^2, x_3^2)$.
   (b) $F(x, y, z) = (xz^2, 0, z^3)$.

6. Demuestre que si R es una región para la cual el teorema de Gauss es válido, entonces el volumen de R es
   $$V(R) = \frac{1}{3} \int_{\partial R} x_1 \, dx_2 \, dx_3 + x_2 \, dx_3 \, dx_1 + x_3 \, dx_1 \, dx_2.$$

7. (a) Use el teorema de Gauss para demostrar que si F es un campo vectorial continuamente diferenciable con divergencia cero en una región R, entonces la integral de F sobre ∂R es cero.
   (b) Escriba un argumento intuitivo basado en la interpretación de la divergencia, para demostrar la afirmación de la parte (a).

8. Sea S el elipsoide
   $$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$
   y sea $D(x, y, z)$ la distancia del origen al plano tangente a S en (x, y, z).

(a) Demuestre que si $F(x, y, z) = (x/a^2, y/b^2, z/c^2)$, entonces $F \cdot \mathbf{n} = D^{-1}$, donde $\mathbf{n}$ es el vector unitario normal a S en (x, y, z) y que apunta hacia afuera.

(b) Demuestre que $\int_S D^{-1}\, d\sigma = \frac{4\pi}{3}\left(\frac{bc}{a} + \frac{ca}{b} + \frac{ab}{c}\right)$.

9. Un campo vectorial $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}^3$ definido en una región R se llama **irrotacional** si rot $F(\mathbf{x}) = 0$ para todo x en R y se llama **incompresible** si div $F(\mathbf{x}) = 0$ para todo x en R. Suponga que F es continuamente diferenciable.

(a) Demuestre que si F es irrotacional, entonces la circulación de F a lo largo de un camino circular suficientemente pequeño en R, es cero.

(b) Demuestre que si F es incompresible, entonces el flujo de F es cero a través de toda esfera suficientemente pequeña y tal que su interior está contenida en R.

6. Los operadores ∇, $\nabla\times$ y $\nabla\cdot$

Para facilitar la aplicación de los teoremas de Gauss y de Stokes es útil extender el uso del símbolo ∇, llamado «del», que hemos usado para denotar al gradiente de una función real. Recordemos que, en términos de la base natural e_1, e_2, e_3 de $\mathcal{R}^3$, tenemos que

$$\nabla f = \frac{\partial f}{\partial x_1}\mathbf{e}_1 + \frac{\partial f}{\partial x_2}\mathbf{e}_2 + \frac{\partial f}{\partial x_3}\mathbf{e}_3. \tag{1}$$

Esta ecuación define ∇ como un operador definido para las funciones reales diferenciables $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}$ y con campos de vectores $\mathcal{R}^3 \xrightarrow{F} \mathcal{R}^3$ como valores. Si escribimos

$$\nabla = \frac{\partial}{\partial x_1}\mathbf{e}_1 + \frac{\partial}{\partial x_2}\mathbf{e}_2 + \frac{\partial}{\partial x_3}\mathbf{e}_3, \tag{2}$$

entonces la Ec. (1) se obtiene aplicando a f ambos lados de la Ec. (2).

Tenemos entonces de manera natural las definiciones siguientes. Si F es un campo vectorial diferenciable dado por

$$F(\mathbf{x}) = F_1(x)\mathbf{e}_1 + F_2(\mathbf{x})\mathbf{e}_2 + F_3(\mathbf{x})\mathbf{e}_3,$$

entonces el operador $\nabla\times$ se define formando el producto cruz formal de ∇ y F, obteniéndose

$$\nabla \times F = \begin{vmatrix} \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_2 & F_3 \end{vmatrix}\mathbf{e}_1 + \begin{vmatrix} \frac{\partial}{\partial x_3} & \frac{\partial}{\partial x_1} \\ F_3 & F_1 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \\ F_1 & F_2 \end{vmatrix}\mathbf{e}_3$$

$$= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}\right)\mathbf{e}_1 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}\right)\mathbf{e}_2 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right)\mathbf{e}_3.$$

Por tanto, $\nabla \times F$ es lo que hemos denotado por rot F. Análogamente, definimos el operador $\nabla\cdot$ tomando el producto punto formal de ∇ y F, obteniéndose

$$\nabla \cdot F = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}. \tag{3}$$

Esta es la función real que hemos denotado por div F. La razón para esta notación es que hace fácil recordar las definiciones siempre que se recuerde la Ec. (2).

Con esta notación, la fórmula de Stokes se escribe en la forma

6.1 $$\int_S (\nabla \times F)\cdot \mathbf{n}\, d\sigma = \int_{\partial S} F \cdot \mathbf{t}\, ds,$$

y la fórmula de Gauss como

6.2 $$\int_R \nabla \cdot F\, dV = \int_{\partial R} F \cdot \mathbf{n}\, d\sigma.$$

Para hacer buen uso de estas fórmulas necesitamos algunas identidades para ∇. En las fórmulas que siguen, f y g son funciones reales diferenciables, F y G son campos vectoriales diferenciables y a y b son constantes.

$$\nabla(af + bg) = a\nabla f + b\nabla g \tag{4}$$

$$\nabla(fg) = f\nabla g + g\nabla f \tag{5}$$

$$\nabla \times (aF + bG) = a\nabla \times F + b\nabla \times G \tag{6}$$

$$\nabla \times (fF) = f\nabla \times F + \nabla f \times F \tag{7}$$

$$\nabla \cdot (aF + bG) = a\nabla \cdot F + b\nabla \cdot G \tag{8}$$

$$\nabla \cdot (fF) = f\nabla \cdot F + \nabla f \cdot F \tag{9}$$

$$\nabla \cdot (F \times G) = (\nabla \times F)\cdot G - F \cdot (\nabla \times G). \tag{10}$$

Cada una de estas fórmulas es una consecuencia inmediata de las definiciones de los operadores. El mismo método de demostración establece que si f y F son dos veces diferenciables, entonces

$$\nabla \cdot (\nabla \times F) = 0, \tag{11}$$

$$\nabla \times (\nabla f) = 0, \tag{12}$$

$$\nabla \cdot \nabla f = \nabla^2 f, \tag{13}$$

donde $\nabla^2 f$ es el laplaciano de f, definido por

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial x^2 f}{\partial x_3^2}.$$

[Las Ecs. (11), (12) y (13) son las mismas de los Ejercicios 2 y 3 de la sección anterior, en los cuales se usó en símbolo Δ para ∇^2.]

Estas fórmulas permiten deducir varios casos especiales de los teoremas de Gauss y de Stokes. Un caso especialmente importante es cuando se supone que el campo vectorial F es un gradiente ∇f o un múltiplo $f\nabla g$. Si hacemos $F = \nabla f$ en la Ec. 6.2, tenemos que

$$\int_R \nabla \cdot \nabla f \, dV = \int_{\partial R} \nabla f \cdot \mathbf{n} \, d\sigma. \tag{14}$$

Pero, de acuerdo con la Ec. (13), $\nabla \cdot \nabla f = \nabla^2 f$ y, de acuerdo con la Ec. 2.1 del Cap. 4, $\nabla f \cdot \mathbf{n} = (\partial/\partial \mathbf{n}) f$. Por tanto,

$$\int_R \nabla^2 f \, dV = \int_S \frac{\partial f}{\partial \mathbf{n}} \, d\sigma. \tag{15}$$

Si en la Ec. 6.2 remplazamos F por $f\nabla g$, usando la Ec. (9) tenemos que

$$\nabla \cdot (f \nabla g) = f \nabla \cdot \nabla g + \nabla f \cdot \nabla g;$$

de modo que de la fórmula de Gauss se tiene que

6.3
$$\int_R f \nabla^2 g \, dV + \int_R \nabla f \cdot \nabla g \, dV = \int_S f \frac{\partial g}{\partial \mathbf{n}} \, d\sigma.$$

Esta igualdad se conoce como la **primera identidad de Green.** Debido a la simetría en el término del medio, si intercambiamos f y g y restamos los términos correspondientes, tenemos la **segunda identidad de Green.**

6.4
$$\int_R (f \nabla^2 g - g \nabla^2 f) \, dV = \int_S \left(f \frac{\partial g}{\partial \mathbf{n}} - g \frac{\partial f}{\partial \mathbf{n}} \right) d\sigma.$$

Ejemplo 1. Sea R una región poligonalmente conexa en $\mathfrak{R}^3$ con una frontera lisa por pedazos S. Si h es una función real definida en R, consideramos la **ecuación de Poisson**

$$\nabla^2 u = h,$$

con la condición de frontera $u(\mathbf{x}) = \varphi(\mathbf{x})$ para $\mathbf{x}$ en S. Suponemos que existe por lo menos una solución $u(\mathbf{x})$ definida en R y que satisface la condición de frontera. Usando la primera identidad de Green, podemos demostrar que *esa solución es única*. Supongamos que existen dos soluciones u_1 y u_2. Entonces la función u definida por $u(\mathbf{x}) = u_1(\mathbf{x}) - u_2(\mathbf{x})$ satisface la **ecuación de Laplace** $\nabla^2 u = 0$ en R con la condición de frontera $u(\mathbf{x}) = 0$ en S. Tomando $f = g = u$ en la Ec. 6.3, se tiene que

$$\int_R u \nabla^2 u \, dV + \int_R |\nabla u|^2 \, dV = \int_S u \frac{\partial u}{\partial \mathbf{n}} \, d\sigma.$$

Pero el primer y el último término son cero, ya que $\nabla^2 u = 0$ en R y $u = 0$ en S. Por tanto, $\int_R |\nabla u|^2 \, dV = 0$ y $\nabla u = 0$ idénticamente en R y en S. Por tanto, u debe ser constante en la región poligonalmente conexa R. Finalmente, u debe ser idénticamente cero, ya que $u(\mathbf{x}) = 0$ para $\mathbf{x}$ en S. Notemos que la

ecuación de Laplace es un caso especial de la de Poisson, obtenida tomando h idénticamente cero, de modo que también hemos demostrado un teorema de unicidad para la ecuación de Laplace.

Ejemplo 2. Como una aplicación de la segunda identidad de Green, consideramos el problema siguiente. Sea D un operador diferencial definido para funciones definidas en una región R de $\mathfrak{R}^3$ que tiene una superficie lisa por pedazos S como frontera. Si denotamos por $\mathfrak{F}$ el espacio vectorial de las funciones continuas definidas en R y S, podemos definir un producto interno en $\mathfrak{F}$ por

$$\langle f, g \rangle = \int_R fg \, dV.$$

Es fácil verificar que esto es un producto interno. A menudo es importante saber que el operador D es simétrico con respecto de este producto interno en el sentido que

$$\langle Df, g \rangle = \langle f, Dg \rangle,$$

cuando Df y Dg están definidas y son continuas. Para la definición análoga para espacios de dimensión finita, sea el Ejercicio 17 de la Sec. 2 del Cap. 5. Para resolver ese problema es necesario imponer a las funciones f en $\mathfrak{F}$ condiciones de frontera apropiadas. Así podemos demostrar que el operador de Laplace ∇^2 es simétrico cuando se lo define en el subespacio $\mathfrak{F}_0$ de $\mathfrak{F}$ formado por las funciones continuas f que satisfacen una condición de frontera de la forma

$$a_1 f(\mathbf{x}) + a_2 \frac{\partial f}{\partial \mathbf{n}}(\mathbf{x}) = 0. \tag{16}$$

En la Ec. (16), a_1 y a_2 son constantes tales que una de ellas es distinta de cero y $\partial/\partial u$ denota la diferenciación con respecto del vector unitario normal a S que apunta hacia afuera.

Supongamos que f y g son dos funciones que satisfacen la Ec. (16) y tales que $\nabla^2 f$ y $\nabla^2 g$ son continuas. Entonces la segunda identidad de Green (Ec. 6.4) muestra que $\langle \nabla^2 f, g \rangle = \langle f, \nabla^2 g \rangle$ si, y solo si,

$$\int_S \left(f \frac{\partial f}{\partial \mathbf{n}} - g \frac{\partial f}{\partial \mathbf{n}} \right) d\sigma = 0. \tag{17}$$

Como f y g satisfacen la Ec. (16), tenemos que

$$g a_1 f + g a_2 \frac{\partial f}{\partial \mathbf{n}} = 0,$$

$$f a_1 g + f a_2 \frac{\partial g}{\partial \mathbf{n}} = 0,$$

y restando estas ecuaciones tenemos que

$$a_2 \left[g(\mathbf{x}) \frac{\partial f}{\partial \mathbf{n}}(\mathbf{x}) - f(\mathbf{x}) \frac{\partial g}{\partial \mathbf{n}}(\mathbf{x}) \right] = 0, \text{ para todo } \mathbf{x} \text{ en } S.$$

Si $a_2 \neq 0$, esto implica que $g(\partial f/\partial \mathfrak{n}) - f(\partial g/\partial \mathfrak{n}) = 0$ idénticamente en S, de modo que la Ec. (17) es satisfecha. Por otra parte, si $a_2 = 0$, entonces a_1 no es cero y la condición de frontera implica que f y g son idénticamente cero en S. Pero entonces la Ec. (17) sigue siendo válida. Por tanto, hemos demostrado que ∇^2 es simétrico en $\mathfrak{F}_0$. Nótese que $\nabla^2 f$ solo está definido para f en $\mathfrak{F}_0$ suficientemente diferenciable.

De la Ec. 6.2 podemos deducir algunas ecuaciones para integrales vectoriales. En el Ejercicio 16 de la Sec. 2 del Cap. 6 se ha dado la definición y algunas propiedades de esta integral. Sea $\mathbf{k}$ un vector constante arbitrario y sea $F(\mathbf{x}) = f(\mathbf{x})\mathbf{k}$, donde f es una función real continuamente diferenciable en una región R. Entonces, como $\nabla \cdot f\mathbf{k} = \nabla f \cdot \mathbf{k}$ (¡verifique!) de la Ec. 6.2, se tiene que

$$\int_R \nabla f \cdot \mathbf{k}\, dV = \int_{\partial R} f\mathbf{k} \cdot \mathbf{n}\, d\sigma.$$

Como $\mathbf{k}$ es constante,

$$\mathbf{k} \cdot \int_R \nabla f\, dV = \mathbf{k} \cdot \int_{\partial R} f\mathbf{n}\, d\sigma,$$

y, como $\mathbf{k}$ es arbitrario,

$$\int_R \nabla f\, dV = \int_{\partial R} f\mathbf{n}\, d\sigma. \tag{18}$$

Análogamente, remplazando F por $\mathbf{k} \times F$ en la Ec. 6.2, donde $\mathbf{k}$ es un vector constante, podemos demostrar que

$$\int_R \nabla \times F\, dV = \int_{\partial R} \mathbf{n} \times F\, d\sigma. \tag{19}$$

La demostración se deja como ejercicio.

Terminamos esta sección con expresiones para el gradiente, divergencia y rotor con respecto de un sistema curvilíneo ortogonal de coordenadas. Por tanto, suponemos que todas las curvas coordenadas se cortan en ángulos rectos, lo cual significa que los vectores tangentes naturales $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ son perpendiculares en cada punto. Si f es una función real o F es un campo vectorial definimos una región D en $\mathfrak{R}^3$, entonces puede ser conveniente sustituir las variables coordenadas curvilíneas u_1, u_2, u_3 usando una transformación,

$$(x_1, x_2, x_3) = T(u_1, u_2, u_3).$$

Esto conduce a una nueva función real $\bar{f}$ definida en la región del espacio u correspondiente a D por

$$f \circ T = \bar{f}.$$

En el caso de un campo vectorial hacemos un cambio adicional y lo expresamos en términos de los vectores unitarios

$$\frac{1}{h_1}\mathbf{c}_1, \frac{1}{h_2}\mathbf{c}_2, \frac{1}{h_3}\mathbf{c}_3,$$

donde $h_i = |\mathbf{c}_i|$. Obtenemos así funciones coordenadas $\bar{F}_1$, $\bar{F}_2$, $\bar{F}_3$ que satisfacen

$$F \circ T = \frac{1}{h_1}\bar{F}_1\mathbf{c}_1 + \frac{1}{h_2}\bar{F}_2\mathbf{c}_2 + \frac{1}{h_3}\bar{F}_3\mathbf{c}_3.$$

Mediante un cálculo bastante largo, usando la regla de la cadena, se obtienen las fórmulas

6.5
$$(\nabla f) \circ T = \frac{1}{h_1^2}\frac{\partial \bar{f}}{\partial u_1}\mathbf{c}_1 + \frac{1}{h_2^2}\frac{\partial \bar{f}}{\partial u_2}\mathbf{c}_2 + \frac{1}{h_3^2}\frac{\partial \bar{f}}{\partial u_3}\mathbf{c}_3.$$

6.6
$$(\nabla \times F) \circ T = \left[\frac{\partial h_2 h_3 \bar{F}_1}{\partial u_1} + \frac{\partial h_1 h_3 \bar{F}_2}{\partial u_2} + \frac{\partial h_1 h_2 \bar{F}_3}{\partial u_3}\right]$$

6.7
$$\begin{aligned}(\nabla \times F) \circ T = \frac{\pm 1}{h_1 h_2 h_3}\Bigg[&\left(\frac{\partial h_3 \bar{F}_3}{\partial u_2} - \frac{\partial h_2 \bar{F}_2}{\partial u_3}\right)\mathbf{c}_1 \\ &+ \left(\frac{\partial h_1 \bar{F}_1}{\partial u_3} - \frac{\partial h_3 \bar{F}_3}{\partial u_1}\right)\mathbf{c}_2 \\ &+ \left(\frac{\partial h_2 \bar{F}_2}{\partial u_1} - \frac{\partial h_1 \bar{F}_1}{\partial u_2}\right)\mathbf{c}_3\Bigg]\end{aligned}$$

En la última fórmula se elige el signo más cuando los $\mathbf{c}_i$ forman un sistema derecho, y el signo menos en caso contrario. Es claro que los $\mathbf{c}$ y los h pueden variar de punto en punto. Las fórmulas *no* son más complicadas por haber supuesto que los $\mathbf{c}$ son perpendiculares. Un resultado es que $h_1h_2h_3$ puede ser interpretado como el volumen de la caja rectangular generada por $\mathbf{c}_1$, $\mathbf{c}_2$ y $\mathbf{c}_3$. En el caso de coordenadas rectangulares en $\mathfrak{R}^3$, la transformación T es la identidad y las fórmulas resultantes se reducen a las definiciones correspondientes en coordenadas rectangulares.

Ejemplo 3. Si introducimos coordenadas esféricas en $\mathfrak{R}^3$, del Ejemplo 4 de la Sec. 5 del Cap. 4, se tiene que

$$h_1 = 1, \qquad h_2 = r, \qquad h_3 = r \operatorname{sen} \varphi.$$

Entonces en el punto (x, y, z) en $\mathfrak{R}^3$ con coordenadas (r, φ, θ) tenemos que

$$\nabla f(x, y, z) = \frac{\partial \bar{f}}{\partial r}(r, \varphi, \theta)\mathbf{c}_1 + \frac{1}{r^2}\frac{\partial \bar{f}}{\partial \varphi}(r, \varphi, \theta)\mathbf{c}_2 + \frac{1}{r^2 \operatorname{sen}^2 \varphi}\frac{\partial \bar{f}}{\partial \theta}(r, \varphi, \theta)\mathbf{c}_3.$$

y

$$\begin{aligned}\nabla \cdot F(x, y, z) = \frac{1}{r^2 \operatorname{sen} \varphi}\Bigg[&\operatorname{sen} \varphi \frac{\partial r^2 \bar{F}_1(r, \varphi, \theta)}{\partial r} + r\frac{\partial \operatorname{sen} \varphi \bar{F}_2(r, \varphi, \theta)}{\partial \varphi} \\ &+ r\frac{\partial \bar{F}_3(r, \varphi, \theta)}{\partial \theta}\Bigg]\end{aligned}$$

Como el laplaciano está definido por $\nabla^2 f = \nabla \cdot \nabla f$, tenemos que

$$\nabla^2 f(x, y, z) = \frac{1}{r^2 \operatorname{sen}^2 \varphi}\left[\operatorname{sen}^2 \varphi \frac{\partial}{\partial r}\left(r^2 \frac{\partial \bar{f}(r, \varphi, \theta)}{\partial r}\right)\right.$$
$$\left. + \operatorname{sen} \varphi \frac{\partial}{\partial \varphi}\left(\operatorname{sen} \varphi \frac{\partial \bar{f}(r, \varphi, \theta)}{\partial \varphi} + \frac{\partial^2 \bar{f}(r, \varphi, \theta)}{\partial \theta^2}\right)\right]$$

Al verificar este cálculo es importante recordar que la Ec. 6.6 está relacionada con las funciones coordenadas con respecto de vectores unitarios en la dirección de los $\mathbf{e}$ y no de los $\mathbf{e}$ mismos.

EJERCICIOS

1. Verifique las identidades (4) a (10).

2. Si $F = (F_1, F_2, F_3)$ es un campo vectorial, defina el operador $F \cdot \nabla$ como

$$F_1 \frac{\partial}{\partial x_1} + F_2 \frac{\partial}{\partial x_2} + F_3 \frac{\partial}{\partial x_3}.$$

(a) Si $F(\mathbf{x}) = \mathbf{x}$ calcule $(\mathbf{x} \cdot \nabla)G(\mathbf{x})$, donde $G(x_1, x_2, x_3) = (x_1^2, x_1, x_2, x_3)$.
(b) Demuestre que, en general, $(\nabla \cdot F)G \neq (F \cdot \nabla)G$.
(c) Si $\mathbf{k}$ es un vector constante, demuestre que

$$\nabla \times (\mathbf{k} \times F) = \mathbf{k}(\nabla \cdot F) - (\mathbf{k} \cdot \nabla)F,$$

donde F es un campo diferenciable de vectores.
(d) Use la parte (c) para demostrar que si F y G son diferenciables, entonces

$$\nabla \times (F \times G) = F(\nabla \cdot G) - (F \cdot \nabla)G - G(\nabla \cdot F) + (G \cdot \nabla)F.$$

3. Demuestre que si $\mathbf{k}$ es un vector constante, entonces

$$\nabla \times \frac{\mathbf{k} \times \mathbf{x}}{|\mathbf{x}|} = \frac{\mathbf{k}}{|\mathbf{x}|} + \frac{\mathbf{k} \cdot \mathbf{x}}{|\mathbf{x}|^3}\,\mathbf{x}.$$

4. Demuestre que

(a) $\nabla\left(\frac{1}{|\mathbf{x}|}\right) = \frac{-\mathbf{x}}{|\mathbf{x}|^3}, \quad \mathbf{x} \neq 0.$

(b) $\nabla^2\left(\frac{1}{|\mathbf{x}|}\right) = 0, \quad \mathbf{x} \neq 0.$

5. Si $T(\mathbf{x})$ es la temperatura de estado estable en un punto $\mathbf{x}$ de un conjunto abierto en R en $\mathfrak{R}^3$, entonces el flujo del gradiente de temperatura a través de una superficie cerrada lisa en R es cero. Use este hecho y la Ec. (15) para demostrar que la función temperatura en un estado estable dos veces continuamente diferenciable es armónica, es decir, $\nabla^2 T \equiv 0$. [*Indicación*: Suponga que $\nabla^2 T(\mathbf{x}_0) > 0$. Demuestre que $\nabla^2 T(\mathbf{x}) > 0$ en alguna bola de centro $\mathbf{x}_0$.]

6. La condición (16) del Ejemplo 2 puede ser generalizada a

$$\phi_1(\mathbf{x}) f(\mathbf{x}) + \phi_2(\mathbf{x}) \frac{\partial f}{\partial \mathbf{n}}(\mathbf{x}) = 0,$$

donde ϕ_1 y ϕ_2 son funciones continuas que satisfacen $\phi_1^2(\mathbf{x}) + \phi_2^2(\mathbf{x}) > 0$. Demuestre que el operador de Laplace ∇^2 es simétrico con esta condición más general.

7. Demuestre que si $\nabla \cdot F$ es idénticamente cero en una bola B en $\mathfrak{R}^3$, entonces el campo vectorial G definido por

$$G(\mathbf{x}) = \int_0^1 [F(t\mathbf{x}) \times (t\mathbf{x})]\, dt$$

satisface la relación $\nabla \times G(\mathbf{x}) = F(\mathbf{x})$ en B. La demostración consiste en los siguientes pasos:
(a) Si $G(\mathbf{x}) = \int_0^1 [F(t\mathbf{x}) \times (t\mathbf{x})]\, dt$, demuestre que

$$\nabla \times G(\mathbf{x}) = \int_0^1 \nabla \times [F(t\mathbf{x}) \times (t\mathbf{x})]\, dt.$$

[*Indicación*: Aplique la regla de Leibnitz al Ejercicio 7(a) de la Sec. 2 del Cap. 6.]
(b) Usando la identidad del Ejercicio 2(d), demuestre que $\nabla \times [F(t\mathbf{x}) \times (t\mathbf{x})] = 2tF(t\mathbf{x}) + t^2(d/dt)F(t\mathbf{x})$.
(c) Demuestre que $\nabla \times G(\mathbf{x}) = F(\mathbf{x})$.

8. Un campo vectorial $\mathfrak{R}^3 \xrightarrow{F} \mathfrak{R}^3$ definido en una región R se llama si, en una vecindad B de cada punto de R, F puede ser representada como el rotor de un campo G_B. El campo F es **incompresible** si la divergencia de F es cero en cada punto de R. Demuestre que un campo continuamente diferenciable de vectores es solenoidal si, y solo si, es incompresible. Para una interpretación de la incompresibilidad, véase el Ejercicio 9 de la sección anterior.

9. Considere la función **potencial newtoniano** $N(\mathbf{x}) = |\mathbf{x}|^{-1}$ y su campo gradiente asociado $\nabla N(\mathbf{x})$. (Véase el Ejercicio 4.) Demuestre que $N(\mathbf{x})$ puede ser interpretado como el trabajo efectuado al mover una partícula desde ∞ a $\mathbf{x}$ a lo largo de un camino liso a través del campo ∇N.

10. Sea R una región acotada en $\mathfrak{R}^3$ y $p(\mathbf{x})$ la densidad en el punto $\mathbf{x}$ en R. Entonces la integral

$$N_p(\mathbf{x}) = \int_R \frac{p(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|}\, dV_y$$

se llama el **potencial newtoniano** de la distribución de materia con densidad p en R.

(a) Demuestre que si p es continua en un conjunto abierto R con frontera lisa, entonces, para $\mathbf{x}$ *fuera* de R,

$$\nabla^2 N_p = 0.$$

En otras palabras, demuestre que para $\mathbf{x}$ fuera de R, $N_p(\mathbf{x})$ es armónica.
(b) Bajo las suposiciones anteriores sobre p y R, demuestre que la integral $N_p(\mathbf{x})$ existe como integral impropia cuando $\mathbf{x}$ está en R.

11. (a) Use las Ecs. 6.5 y 6.6 para demostrar que el laplaciano tiene la forma siguiente en coordenadas cilíndricas

$$\nabla^2 f(x, y, z) = \frac{1}{r^2}\left[r \frac{\partial}{\partial r}\left(r \frac{\partial f(r, \theta, z)}{\partial r}\right) + \frac{\partial^2 f(r, \theta, z)}{\partial \theta^2} + r^2 \frac{\partial^2 f(r, \theta, z)}{\partial z^2}\right].$$

(b) Demuestre que si $f(x, y, z)$ es independiente de z, entonces

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f(r,\theta)}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f(r, \theta)}{\partial \theta^2} + \frac{1}{r}\frac{\partial f(r, \theta)}{\partial r}.$$

en coordenadas cilíndricas. Compare con el Ejercicio 13 de la Sec. 5 del Cap. 4.

(c) Demuestre que si $f(x, y, z)$ puede ser expresada como una función $\bar{f}(\sqrt{x^2 + y^2})$, entonces

$$\nabla^2 f(x, y, z) = \frac{\partial^2 \bar{f}(r)}{\partial r^2} + \frac{1}{r}\frac{\partial f(r)}{\partial r}.$$

12. Demuestre que si $f(x, y, z)$ puede ser descrita como una función $\bar{f}(\sqrt{x^2 + y^2 + z^2})$, entonces

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f(r)}{\partial r^2} + \frac{2}{r}\frac{\partial f(r)}{\partial r}.$$

7. Formas diferenciales

Al definir la integral de línea en el Cap. 3, notamos que, en el caso tridimensional, ella puede ser abreviada en la forma

$$\int_\gamma F_1\, dx_1 + F_2\, dx_2 + F_3\, dx_3$$

En esta sección queremos demostrar que el integrando $F_1\, dx_1 + F_2\, dx_2 + F_3\, dx_3$ tiene una interpretación que conduce a una nueva manera de considerar la integral de línea. Esto nos conducirá a una definición natural de la integral de superficie.

Denotamos por dx_k la función que asocia a cada vector $\mathbf{a}$ de $\mathcal{R}^n$ su k-ésima coordenada. Así, si $\mathbf{a} = (a_1, \ldots, a_k, \ldots, a_n)$ está en $\mathcal{R}^n$, entonces $dx_k(\mathbf{a}) = a_k$. En particular, si $\mathbf{a} = (-1, 1, 3)$, entonces $dx_1(\mathbf{a}) = -1$, $dx_2(\mathbf{a}) = 1$ y $dx_3(\mathbf{a}) = 3$. Geométricamente, $dx_k(\mathbf{a})$ es el largo, con signo apropiado, de la proyección de $\mathbf{a}$ en el k-ésimo eje de coordenadas, tal como lo muestra la Fig. 31. Las com-

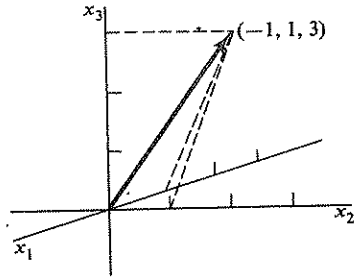


Figura 31

binaciones lineales con coeficientes constantes, de las funciones dx_k, definen nuevas funciones

$$c_1\, dx_1 + \ldots + c_n\, dx_n.$$

Es más, si $F_1, \ldots, F_n$ son funciones reales definidas en una región D de $\mathcal{R}^n$ para cada $\mathbf{x}$ en D podemos formar la combinación lineal

$$\omega_{\mathbf{x}} = F_1(\mathbf{x})\, dx_1 + \ldots + F_n(\mathbf{x})\, dx_n, \tag{1}$$

donde $\omega_{\mathbf{x}}$ está definida por

$$\omega_{\mathbf{x}}(\mathbf{a}) = F_1(\mathbf{x})\, dx_1(\mathbf{a}) + \ldots + F_n(\mathbf{x})\, dx_n(\mathbf{a}). \tag{2}$$

para todo $\mathbf{a}$ en $\mathcal{R}^n$. Por ejemplo, si $\omega_{(x,y)} = x^2\, dx + y^2\, dy$ en $\mathcal{R}^2$, entonces $\omega_{(x,y)}(a, b) = ax^2 + by^2$ y $\omega_{(-1,3)}(a, b) = a + 9b$. Una función $\omega_{\mathbf{x}}$ definida como en la Ec. (2) se llama una **1-forma diferenciable** o simplemente **1-forma.**

Ejemplo 1. Sea $\mathcal{R}^3 \xrightarrow{f} \mathcal{R}$ una función diferenciable en una región D en $\mathcal{R}^3$. Entonces la diferencial de f en $\mathbf{x}$, $d_{\mathbf{x}}f$, es una 1-forma en D, ya que si $\mathbf{a} = (a_1\ a_2,\ a_3)$, entonces

$$\begin{aligned} d_{\mathbf{x}}f(\mathbf{a}) &= \frac{\partial f}{\partial x_1}(\mathbf{x})a_1 + \frac{\partial f}{\partial x_2}(\mathbf{x})a_2 + \frac{\partial f}{\partial x_3}(\mathbf{x})a_3 \\ &= \frac{\partial f}{\partial x_1}(\mathbf{x})\, dx_1(\mathbf{a}) + \frac{\partial f}{\partial x_2}(\mathbf{x})\, dx_2(\mathbf{a}) + \frac{\partial f}{\partial x_3}(\mathbf{x})\, dx_3(\mathbf{a}). \end{aligned}$$

Por tanto, las funciones coeficientes $F_k(\mathbf{x})$ de la Ec. (1) tienen la forma $F_k(\mathbf{x}) = (\partial f/\partial x_k)(\mathbf{x})$. Sin embargo, no toda 1-forma es la diferencial de una función. Para comprobar esto, recordemos que si f es suficientemente diferenciable, entonces $\partial^2 f/(\partial x_2\, \partial x_1) = \partial^2 f/(\partial x_1, \partial x_2)$, pero en general, $\partial F_1/\partial x_2 \neq \partial F_2/\partial x_1$, a no ser que F_1 y F_2 estén relacionadas de esta manera, como sucede cuando $\nabla f = F$ para algún f.

Ejemplo 2. Si $\omega_{\mathbf{x}}$ es una 1-forma definida en una región D de $\mathcal{R}^3$ y γ es una curva diferenciable en D definida por $\mathcal{R} \xrightarrow{g} \mathcal{R}^3$ para $a \leq t \leq b$, entonces en cada punto $\mathbf{x} = g(t)$ de γ podemos aplicar la función lineal $\omega_{g(t)}$ al vector tangente en $\mathbf{x}$, $g'(t)$. El resultado es un número real que expresamos en la forma

$$\begin{aligned} \omega_{g(t)}(g'(t)) &= F_1(g(t))\, dx_1(g'(t)) + F_2(g(t))\, dx_2(g'(t)) + F_3(g(t))\, dx_3(g'(t)) \\ &= F_1(g(t))g_1'(t) + F_2(g(t))g_2'(t) + F_3(g(t))g_3'(t) \\ &= F(g(t)) \cdot g'(t). \end{aligned}$$

Si escribimos

$$\int_a^b \omega_{g(t)}(g'(t))\, dt = \int_a^b F(g(t)) \cdot g'(t)\, dt,$$

vemos que la integral de la derecha es la integral de línea del campo F a lo largo de γ.

El ejemplo anterior sugiere una definición natural de la integral de una 1-forma a lo largo de una curva γ. Si $\omega_{\mathbf{x}}$ está definida en una región D en $\mathcal{R}^n$ por la Ec. (1) y γ es una curva en D parametrizada por $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ para $a \leq t \leq b$, entonces podemos hacer dos cosas. Podemos usar las funciones coeficientes $F_1, \ldots, F_n$ de $\omega_{\mathbf{x}}$ y formar un campo vectorial F en D y definir

$$\int_\gamma \omega_{\mathbf{x}} = \int_a^b F(g(t)) \cdot g'(t)\, dt,$$

o podemos formar una partición P del intervalo $a \leq t \leq b$ en puntos $a = t_0 < t_1 < \ldots < t_k = b$ y definir

$$\int_\gamma \omega_{\mathbf{x}} = \lim_{m(P)\to 0} \sum_{k=1}^{K} \omega_{g(t_k)}(g'(t_k))(t_k - t_{k-1}).$$

Es claro que las dos fórmulas dan la misma definición para $\int_\gamma \omega_{\mathbf{x}}$, la *integral de la 1-forma $\omega_{\mathbf{x}}$ a lo largo de γ*.

A continuación definimos un producto de 1-formas que es distinto de la multiplicación punto a punto de funciones. Primero definimos el producto de las 1-formas básicas dx_1, dx_2, dx_3 en $\mathcal{R}^3$. El producto $dx_1 \wedge dx_2$ se define de modo que se obtenga una función definida para pares ordenados de vectores en $\mathcal{R}^3$. Geométricamente, $dx_1 \wedge dx_2(\mathbf{a}, \mathbf{b})$ será el área del paralelogramo generado por las proyecciones de $\mathbf{a}$ y $\mathbf{b}$ en el plano x_1x_2. El signo del área estará determinado de tal manera que si las proyecciones de $\mathbf{a}$ y $\mathbf{b}$ tienen la misma orientación positiva de los ejes x_1 y x_2, entonces el área es positiva. El área será negativa si estas orientaciones son opuestas. La Fig. 32 muestra un caso particular. Así, si $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$, entonces

$$dx_1 \wedge dx_2\,(\mathbf{a}, \mathbf{b}) = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1b_2 - a_2b_1,$$

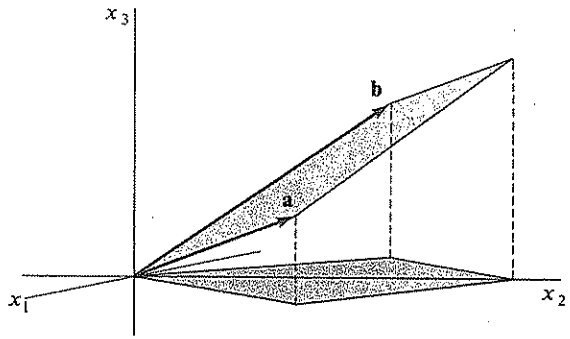


Figura 32

y el determinante asigna el signo correcto al área. Podemos usar las 1-formas básicas dx_1, dx_2, dx_3 para expresar esta última ecuación en la forma

$$dx_1 \wedge dx_2\,(\mathbf{a}, \mathbf{b}) = \det \begin{pmatrix} dx_1\,(\mathbf{a}) & dx_1\,(\mathbf{b}) \\ dx_2\,(\mathbf{a}) & dx_2\,(\mathbf{b}) \end{pmatrix}.$$

Las definiciones de todos los productos de dx_1, dx_2 y dx_3 en cualquier orden están dadas por la ecuación:

$$dx_i \wedge dx_j\,(\mathbf{a}, \mathbf{b}) = \det \begin{pmatrix} dx_i\,(\mathbf{a}) & dx_i\,(\mathbf{b}) \\ dx_j\,(\mathbf{a}) & dx_j\,(\mathbf{b}) \end{pmatrix}, \tag{3}$$

con una interpretación geométrica análoga en cada caso. Por ejemplo,

$$dx_2 \wedge dx_3\,(\mathbf{a}, \mathbf{b}) = \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix},$$

es el área con signo de la proyección del paralelogramo $\mathbf{a}$, $\mathbf{b}$ en el plano x_2, x_3. Como consecuencia de la Ec. (3) y de las propiedades de los determinantes, tenemos las relaciones siguientes:

7.1 $$dx_i \wedge dx_j = -dx_j \wedge dx_i.$$

7.2 $$dx_i \wedge dx_i = 0.$$

La primera ecuación es válida debido a que el intercambio de dos filas adyacentes de un determinante, cambia su signo y la segunda puesto que un determinante con dos filas iguales es cero. Análogamente, si cambiamos columnas, tenemos que

7.3 $$dx_i \wedge dx_j\,(\mathbf{b}, \mathbf{a}) = -dx_i \wedge dx_j\,(\mathbf{a}, \mathbf{b}).$$

De las Ecs. 7.1 y 7.2 es claro que la combinación lineal más general de las funciones $dx_i \wedge dx_j$ puede ser expresada en la forma

$$c_1\, dx_2 \wedge dx_3 + c_2\, dx_3 \wedge dx_1 + c_3\, dx_1 \wedge dx_2.$$

Además, si $F = (F_1, F_2, F_3)$ es un campo vectorial en una región D en $\mathcal{R}^3$, para cada $\mathbf{x}$ en D, podemos definir la función

$$\tau_{\mathbf{x}} = F_1(\mathbf{x})\, dx_2 \wedge dx_3 + F_2(\mathbf{x})\, dx_3 \wedge dx_1 + F_3(\mathbf{x})\, dx_1 \wedge dx_2$$

que tiene como dominio al conjunto de todos los pares ordenados $(\mathbf{a}, \mathbf{b})$ de vectores en $\mathcal{R}^3$. La función $\tau_{\mathbf{x}}$ se llama una **2-forma diferencial**, o **2-forma**.

Ejemplo 3. La 2-forma

$$\tau = 2\, dx_2 \wedge dx_3 + dx_3 \wedge dx_1 + 5\, dx_1 \wedge dx_2$$

es la misma función en cada punto de $\mathcal{R}^3$, ya que sus coeficientes son constantes. (Hemos escrito τ en lugar de $\tau_{\mathbf{x}}$ como lo haremos toda vez que no sea ne-

cesario hacer mención explícita de la variable $\mathbf{x}$.) Sean $\mathbf{a} = (1, 2, 3)$ y $\mathbf{b} = (0, 1, 1)$, entonces

$$\tau(\mathbf{a}, \mathbf{b}) = 2 \det \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} + \det \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} + 5 \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = 2.$$

Recordemos que, como es fácil verificar directamente, el vector $\mathbf{a} \times \mathbf{b}$ con coordenadas

$$\det \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}, \quad \det \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}, \quad \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

es perpendicular a (1, 2, 3) y a (0, 1, 1) y que su largo es igual al área del paralelogramo P generado por $\mathbf{a}$ y $\mathbf{b}$. Por tanto, $\tau(\mathbf{a}, \mathbf{b}) = (2, 1, 5) \cdot (\mathbf{a} \times \mathbf{b})$ es igual a la coordenada de (2, 1, 5) en la dirección perpendicular a P, multiplicada por el área de P. Si interpretamos (2, 1, 5) como el vector velocidad de flujo de un fluido en el espacio, entonces $\tau(a, b)$ será el flujo total a través de P en la unidad de tiempo. La Fig. 33 muestra un ejemplo de un flujo de este tipo. Para un fluido F que se mueve con velocidad constante a través de una superficie S, el **flujo** está definido como la coordenada normal F_n de F, multiplicada por el área de S, es decir, por $(F \cdot \mathbf{n})\sigma(S)$.

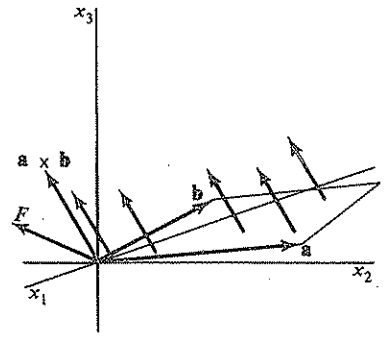


Figura 33

Podemos resumir lo que hemos hecho hasta ahora notando que hemos definido una multiplicación, llamada **producto exterior** de las 1-formas básicas dx_1, dx_2, etc. Los productos resultantes, denotados por $dx_i \wedge dx_j$ son 2-formas *básicas* en el sentido de que toda 2-forma es una combinación lineal de ellas. En $\mathfrak{R}^2$ hay una sola 2-forma básica, $dx_1 \wedge dx_2$, mientras que en $\mathfrak{R}^3$ hay tres. (¿Cuántas hay en $\mathfrak{R}^1$?) Podemos definir el producto exterior de dos 1-formas como la 2-forma obtenida multiplicando las 1-formas como si fueran polinomios en las variables dx_1, dx_2, etc., usando después las reglas 7.1 y 7.2 para hacer simplificaciones. En la práctica, a menudo se omiten las cuñas en la notación.

Ejemplo 4. Para evitar escribir subíndices, denotaremos por dx, dy, dz las 1-formas básicas a pesar de que esta notación es menos conveniente para formular la definición general. Tenemos que

$$\begin{aligned}(x\,dx + y^2\,dy) \wedge (dx + x\,dy) &= x\,dx \wedge dx + y^2\,dy \wedge dx + x^2\,dx \wedge dy + xy^2\,dy \wedge dy \\ &= 0 - y^2\,dx \wedge dy + x^2\,dx \wedge dy + 0 \\ &= (x^2 - y^2)\,dx \wedge dy.\end{aligned}$$

En tres variables tenemos el ejemplo

$$\begin{aligned}(dx + dy + dz) \wedge (x\,dx + z\,dy) &= x\,dx \wedge dx + x\,dy \wedge dx + x\,dz \wedge dx \\ &\quad + z\,dx \wedge dy + z\,dy \wedge dy + z\,dz \wedge dy \\ &= -z\,dy \wedge dz + x\,dz \wedge dx + (z - x)\,dx \wedge dy.\end{aligned}$$

Las 3-formas surgen cuando se trata de definir el producto de una 2-forma y una 1-forma. El significado de las 3-forma básica $dx_1 \wedge dx_2 \wedge dx_3$ es la de una función que asocia un volumen con signo. Así, si $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ y $\mathbf{c} = (c_1, c_2, c_3)$, definimos

$$dx_1 \wedge dx_2 \wedge dx_3(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix},$$

que es el volumen tridimensional orientado del paralelepípedo generado por $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$. Las formas de dimensiones superiores son definidas en el Ejercicio 11.

Una forma diferencial de dimensión no específica se llamará una p-forma y será denotada por ω^p, σ^p, etc. Como lo hiciéramos para 1-formas, definiremos la integral de una p-forma en $\mathfrak{R}^n$ sobre la imagen S por una función continuamente diferenciable g de un conjunto en $\mathfrak{R}^p$. Para simplificar la discusión, supondremos que $\mathfrak{R}^p \xrightarrow{g} \mathfrak{R}^n$ es diferenciable en un rectángulo y acotado R en $\mathfrak{R}^p$. Si ω^p es una p-forma con funciones coeficientes definidas en la imagen $S = g(R)$ en $\mathfrak{R}^n$, podemos aplicar ω^p en un punto $g(\mathbf{u})$ a los vectores derivadas $(\partial g/\partial u_k)(\mathbf{u})$ para $k = 1, \ldots, p$. El resultado tiene la forma

$$\omega^p_{g(\mathbf{u})}\left(\frac{\partial g}{\partial u_1}(\mathbf{u}), \ldots, \frac{\partial g}{\partial u_p}(\mathbf{u})\right). \tag{4}$$

Además, si G es un reticulado en R definido por los puntos $\mathbf{u}_1, \ldots, \mathbf{u}_N$ y con rectángulos $R_1, \ldots, R_N$, podemos formar la suma

$$\sum_{k=1}^{N} \omega^p_{g(\mathbf{u}_k)}\left(\frac{\partial g}{\partial u_1}(\mathbf{u}_k), \ldots, \frac{\partial g}{\partial u_p}(\mathbf{u}_k)\right) V(R_k),$$

donde $V(R_k)$ es el volumen p-dimensional de R_k. Definimos la integral de ω^p sobre S por

$$\int_S \omega^p = \lim_{m(G)\to 0} \sum_{k=1}^{N} \omega^p_{g(u_k)}\left(\frac{\partial g}{\partial u_1}(\mathbf{u}_k), \ldots, \frac{\partial g}{\partial u_p}(\mathbf{u}_k)\right) V(R_k),$$

siempre que este límite exista en el sentido de las sumas de Riemann. Si g es continuamente diferenciable y ω^p tiene funciones coeficientes continuas, entonces la función de $\mathbf{u}$ definida por la Ec. (4) será continua en R y $\int_S \omega^p$ existe y es igual a la integral de Riemann de esa función. Por tanto,

$$\int_S \omega^p = \int_R \omega^p_g\left(\frac{\partial g}{\partial u_1}, \ldots, \frac{\partial g}{\partial u_p}\right) dV.$$

Ejemplo 5. Si

$$\omega^2_{\mathbf{x}} = F_1(\mathbf{x})\, dx_2 \wedge dx_3 + F_2(\mathbf{x})\, dx_3 \wedge dx_1 + F_3(\mathbf{x})\, dx_1 \wedge dx_2$$

tiene coeficientes continuos en $\mathcal{R}^3$ y $\mathcal{R}^2 \xrightarrow{g} \mathcal{R}^3$ es continuamente diferenciable en R con funciones coordenadas g_1, g_2, g_3, entonces

$$\omega_g\left(\frac{\partial g}{\partial u_1}, \frac{\partial g}{\partial u_2}\right) = F_1 \circ g \frac{\partial(g_2, g_3)}{\partial(u_1, u_2)} + F_2 \circ g \frac{\partial(g_3, g_1)}{\partial(u_1, u_2)} + F_3 \circ g \frac{\partial(g_1, g_2)}{\partial(u_1, u_2)}.$$

Por tanto, en este caso,

$$\begin{aligned}\int_S \omega^p &= \int_S F_1\, dx_2 \wedge dx_3 + F_2\, dx_3 \wedge dx_1 + F_3\, dx_1 \wedge dx_2 \\ &= \int_R \left[F_1 \circ g \frac{\partial(g_2, g_3)}{\partial(u_1, u_2)} + F_2 \circ g \frac{\partial(g_3, g_1)}{\partial(u_1, u_2)} + F_3 \circ g \frac{\partial(g_1, g_2)}{\partial(u_1, u_2)}\right] du_1\, du_2.\end{aligned}$$

Salvo la condición de diferenciabilidad en S que hemos introducido por conveniencia, esta definición es igual a la de integral de superficie definida en la Sec. 3.

EJERCICIOS

1. Encuentre el valor de las formas diferenciales en los vectores o conjuntos ordenados de vectores correspondientes que se indican a continuación.
   (a) $dx_1 + 2dx_2$; $(1, 1)$.
   (b) $3dx - dy + dz$; $(1, -1, 0)$.
   (c) $dx_2 + dx_3$; $(1, 3, -5)$.
   (d) $dx_1 + 2\,dx_2 + \ldots + n\,dx_n$; $(1, -1, 1, -1, \ldots)$.
   (e) $2dx_1 \wedge dx_2$; $((1, 1), (1, -1))$.
   (f) $dy \wedge dy + 2dx \wedge dz$; $((1, 2, 1), (-1, 2, 3))$.
   (g) $dx \wedge dy \wedge dz$; $((1, 1, 1), (1, 2, 1), (2, 1, 2))$.
   (h) $dx_1 \wedge dx_2$; $((-2, -3, 0), (2, 0, 2))$.

2. Efectúe las multiplicaciones indicadas y simplifique los resultados en
   (a) $(dx_1 + dx_2) \wedge (dx_1 - dx_2)$.
   (b) $(2dx + 3dy - 2dz) \wedge dx$.

   (c) $(dx + dy) \wedge (dx + dy)$.
   (d) $(x^2\,dx + z^2\,dz) \wedge (dx - 2dy)$.
   (e) $(\operatorname{sen} z\,dx + \cos x\,dy) \wedge (dx + dz)$.

3. Calcule $\int_\gamma y\,dx + x\,dy$, donde γ está definida por $g(t) = (\cos t, \operatorname{sen} t)$, $0 \le t \le \pi/2$.
4. Calcule $\int_\gamma x_1\,dx_1 + x_2\,dx_2 + x_3\,dx_3$, donde γ está dada por
   (a) $g(t) = (-t, t^2, t)$ para $-1 \le t \le 1$.
   (b) $h(t) = (t, t, t)$ para $0 \le t \le 1$.
   (c) $k(t) = (t^2, t^2, t^2)$ para $-1 \le t \le 1$.

5. Calcule $\int_\gamma \omega$, donde $\omega = x_1\,dx_1 + x_2^2\,dx_2 + \ldots x_n^n\,dx_n$ y γ está definida por $g(t) = (t, t, t, \ldots)$ para $0 \le t \le 1$.

6. Sea ω y $\bar{\omega}$ 1-formas definidas en una región D de $\mathcal{R}^n$ y f y g funciones reales definidas en D.
   (a) Demuestre que $f\omega + g\bar{\omega}$ define una 1-forma en D.
   (b) Demuestre que si ω, ν y μ son 1-formas en D, entonces
   $$(f\omega + g\nu) \wedge \mu = f\omega \wedge \mu + g\nu \wedge \mu.$$

7. Demuestre que si ω y $\bar{\omega}$ son 1-formas y γ_1 y γ_2 son curvas en las cuales ω y $\bar{\omega}$ son integrales entonces
   $$\int_{\gamma_1} (a\omega + b\bar{\omega}) = a\int_{\gamma_1} \omega + b\int_{\gamma_1} \bar{\omega},$$
   donde a y b son constantes y
   $$\int_{\gamma_1 \cup \gamma_2} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega.$$

8. Sea f una función real continuamente diferenciable definida en una región D de $\mathcal{R}^n$ y γ una curva continuamente diferenciable en D con parametrización $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ para $a \le t \le b$. Demuestre que si $g(a) = \mathbf{a}$ y $g(b) = \mathbf{b}$, entonces
   $$\int_\gamma d_{\mathbf{x}} f = f(\mathbf{b}) - f(\mathbf{a}).$$

9. (a) Demuestre que si $\omega_{\mathbf{x}}$ es una 1-forma en una región D de $\mathcal{R}^n$, entonces para cada $\mathbf{x}_0$ fijo en D, $\omega_{\mathbf{x}_0}$ es una función lineal real en todo $\mathcal{R}^n$.
   (b) Demuestre la conversa de la parte (a), a saber, que si $\omega_{\mathbf{x}_0}$ es una función lineal real definida en todo $\mathcal{R}^n$, entonces $\omega_{\mathbf{x}_0}(\mathbf{a}) = \sum_{k=1}^{n} c_k(\mathbf{x}_0)\, dx_k\,(\mathbf{a})$, para todo $\mathbf{a}$ en $\mathcal{R}^n$, donde $c_k(\mathbf{x}_0)$ son números reales.

10. (a) Sea τ una función real definida para pares $(\mathbf{a}, \mathbf{b})$ de vectores en $\mathcal{R}^3$ tal que $\tau(\mathbf{a}, \mathbf{b}) = -\tau(\mathbf{b}, \mathbf{a})$ y tal que τ es lineal en $\mathbf{a}$ y en $\mathbf{b}$. Demuestre que existe un vector $\mathbf{c}_\tau$ en $\mathcal{R}^3$ tal que $\tau(\mathbf{a}, \mathbf{b}) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c}_\tau)$ para todo $\mathbf{a}$, $\mathbf{b}$ en $\mathcal{R}^3$. [*Indicación*: Demuestre primero que el resultado es verdadero si $\mathbf{a}$ y $\mathbf{b}$ son parte de una base de $\mathcal{R}^3$.]
    (b) Use la parte (a) para demostrar que si $\tau(\mathbf{a}, \mathbf{b}) = -\tau(\mathbf{b}, \mathbf{a})$ y τ es bilineal, entonces τ es una 2-forma en $\mathcal{R}^3$.

11. Si $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p)$ es una p-pla de vectores en $\mathcal{R}^n$, donde $p \geq 1$, defina por

$$dx_{k_1} \wedge dx_{k_2} \wedge \ldots \wedge dx_{k_p}(\mathbf{a}_1, \ldots, \mathbf{a}_p) = \det\left(dx_{k_i}(\mathbf{a}_j)\right)_{\substack{i=1,\ldots,p \\ j=1,\ldots,p}}$$

las p-formas básicas en $\mathcal{R}^n$. Una p-forma es una combinación lineal de p-formas básicas.

(a) Calcule $dx_2 \wedge dx_3 \wedge dx_4 + 2\,dx_1 \wedge dx_2 \wedge dx_4$ $((1, -1, 0, 2), (-1, 1, 1, 1), (0, 1, 2, 0))$.

(b) Demuestre que si se intercambian dos factores adyacentes en una p-forma básica, el signo de la forma cambia.

(c) Demuestre que una p-forma básica con un factor repetido es cero.

(d) Demuestre que la p-forma general puede ser escrita en la forma

$$\omega^p = \sum_{i_1 < \ldots < i_p} f_{i_1,\ldots,i_p}\, dx_{i_1} \wedge \ldots \wedge dx_{i_p},$$

donde $1 \leq i_k \leq n$ para $k = 1, \ldots, p$.

(e) Demuestre que si $p > n$, entonces ω^p es idénticamente cero.

(f) Demuestre que hay $\binom{n}{p}$ términos en la p-forma de la parte (d).

12. Si ω^p y ω^q son una p-forma y una q-forma en $\mathcal{R}^n$ con

$$\omega^p = \sum_{i_1 < \ldots < i_p} f_{i_1,\ldots,i_p}\, dx_{i_1} \wedge \ldots \wedge dx_{i_p}$$

y

$$\omega^p = \sum_{i_1 < \ldots < i_q} g_{i_1,\ldots,i_q}\, dx_{i_1} \wedge \ldots \wedge dx_{i_q},$$

defina su producto exterior $\omega^p \wedge \omega^q$ por

$$\omega^p \wedge \omega^q = \sum f_{i_1 \ldots i_p} g_{i_1 \ldots i_q}\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_q}.$$

(a) Demuestre que si $\omega^p \wedge \omega^q$ es reducida a su forma estándar usando las Ecs. 7.1 y 7.2, entonces se tienen $\binom{n}{p+q}$ términos.

(b) Demuestre que $\omega^p \wedge \omega^q = (-1)^{pq} \omega^q \wedge \omega^p$.

13. Demuestre que la definición de la integral de una p-forma coincide con la definición dada para una 1-forma si $p = 1$.

14. Calcule $\int_S \omega^2$ si $\omega^2 = dx \wedge dy + dx \wedge dz$ y S es la imagen por $g(u, v) = (u \cos v, u \operatorname{sen} v, v)$ de $0 \leq u \leq 1$, $0 \leq v \leq \pi/2$.

15. (a) Si ω^3 es la 3-forma $f\,dx_1 \wedge dx_2 \wedge dx_3$ en $\mathcal{R}^3$ y $\mathcal{R}^3 \xrightarrow{g} \mathcal{R}^3$ es diferenciable, demuestre que

$$\omega^3_g\left(\frac{\partial g}{\partial u_1}, \frac{\partial g}{\partial u_2}, \frac{\partial g}{\partial u_3}\right) = f \circ g\, \frac{\partial(g_1, g_2, g_3)}{\partial(u_1, u_2, u_3)}.$$

(b) Calcule la integral $\int_S dx \wedge dy \wedge dz$, donde S es la imagen por $g(u, v, w) = (u^2, v^2, w^2)$ de $0 \leq u \leq 1$, $0 \leq v \leq 1$, $0 \leq w \leq 1$.

8. La derivada exterior

El teorema fundamental del cálculo dice que si $(d/dx)f$ es integrable en $[a, b]$, entonces

$$\int_a^b \frac{df}{dx}\, dx = f(b) - f(a). \tag{1}$$

Las fórmulas de Stokes y Gauss

$$\int_S \operatorname{rot} F \cdot dS = \int_{\partial S} F \cdot \mathbf{t}\, ds \tag{2}$$

$$\int_R \operatorname{div} F\, dV = \int_{\partial R} F \cdot \mathbf{n}\, d\sigma, \tag{3}$$

son análogas, puesto que ellas expresan la integral de un tipo de derivada de una función en términos de la función en un conjunto de dimensión menor. Mediante las formas diferenciales definiremos la derivación exterior de tal manera de unificar las fórmulas anteriores.

La operación de **derivación exterior** se define inductivamente para las formas diferenciales como sigue. Sea f una función real diferenciable definida en $\mathcal{R}^n$. Entonces, por definición,

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \ldots + \frac{\partial f}{\partial x_n}\, dx_n.$$

Por tanto, la derivada exterior de f es la 1-forma cuyo valor en cada punto del dominio de f es lo que hemos llamado la diferencial de f en ese punto. Para continuar, si

$$\omega^1 = f_1\, dx_1 + \ldots + f_n\, dx_n$$

es una 1-forma con coeficientes diferenciables, entonces definimos

$$d\omega^1 = (df_1) \wedge dx_1 + \ldots + (df_n) \wedge dx_n.$$

Por tanto, $d\omega^1$ es una 2-forma. En general, si ω^p es una p-forma, entonces $d\omega^p$ es la $(p + 1)$-forma obtenida remplazando cada una de las funciones coeficientes de ω^p por la 1-forma que es su derivada exterior. Para hacer consistente la terminología podemos considerar que las funciones reales son 0-formas.

Ejemplo 1. Si $f(x_1, x_2) = x_1^2 + x_2^2$, entonces df está dada por

$$d(x_1^2 + x_2^3) = 2x_1\, dx_1 + 3x_2^2\, dx_2.$$

Si $\omega^1_{(x_1, x_2)} = x_1 x_2\, dx_1 + (x_1^2 + x_2^2)\, dx_2$, entonces $d\omega^1$ está dada por

$$\begin{aligned} d(x_1 x_2\, dx_1 + (x_1^2 + x_2^2)\, dx_2) &= (x_2\, dx_1 + x_1\, dx_2) \wedge dx_1 + (2x_1\, dx_1 + 2x_2\, dx_2) \wedge dx_2 \\ &= x_1\, dx_1 \wedge dx_2. \end{aligned}$$

Si $\omega^2_{(x,y,z)} = xz\,dx \wedge dy + y^2 z\,dx \wedge dz$, entonces $d\omega^2$ está dada por

$$\begin{aligned} d(xz\,dx \wedge dy + y^2 z\,dx \wedge dz) \\ &= (z\,dx + x\,dz) \wedge dx \wedge dy + (2yz\,dy + y^2\,dz) \wedge dx \wedge dz \\ &= (x - 2yz)\,dx \wedge dy \wedge dz. \end{aligned}$$

Usando la derivada exterior podemos formular el teorema general de Stokes en la forma

8.1
$$\int_B d\omega^p = \int_{\partial B} \omega^p,$$

donde B es de dimensión $p + 1$ y ∂B es su frontera de dimensión p. Interpretaremos esta fórmula en varios casos especiales.

Si ω^1 es una 1-forma en $\mathfrak{R}^2$, entonces podemos expresar ω en la forma

$$\omega^1 = F_1\,dx_1 + F_2\,dx_2,$$

y entonces

$$\begin{aligned} d\omega^1 &= \left(\frac{\partial F_1}{\partial x_1}\,dx_1 + \frac{\partial F_1}{\partial x_2}\,dx_2\right) \wedge dx_1 + \left(\frac{\partial F_2}{\partial x_1}\,dx_1 + \frac{\partial F_2}{\partial x_2}\,dx_2\right) \wedge dx_2 \\ &= \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1 \wedge dx_2. \end{aligned} \tag{4}$$

Sustituyendo $d\omega^1$ y ω^1 en la Ec. 8.1, se tiene

$$\int_B \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1 \wedge dx_1 = \int_{\partial B} F_1\,dx_1 + F_2\,dx_2. \tag{5}$$

Esto es casi la fórmula de Green de la Sec. 1, en el caso en que B es un conjunto adecuado en $\mathfrak{R}^2$ y ∂B es la curva orientada en la dirección contraria a la de los punteros del reloj. Sin embargo, la integral de la izquierda de la Ec. (5) ha sido definida en la sección anterior como una integral sobre un conjunto parametrizado, mientras que la integral

$$\int_D \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_2}{\partial x_2}\right) dx_1\,dx_2 = \oint_\gamma F_1\,dx_1 + F_2\,dx_2,$$

de la fórmula de Green es una integral sobre un conjunto sin parametrización. La diferencia entre las dos es que la parametrización puede cubrir más de una vez por su parametrización, mientras que la integral sobre el conjunto D cubre una sola vez cada parte del conjunto. Véase Ejercicio 9.

Si ω^1 es una 1-forma en $\mathfrak{R}^3$ dada por

$$\omega^1 = F_1\,dx_1 + F_2\,dx_2 + F_3\,dx_3,$$

entonces mediante un cálculo como el de la Ec. (4), se tiene que

$$\begin{aligned} d\omega^1 = &\left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}\right) dx_2 \wedge dx_3 \\ &+ \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}\right) dx_3 \wedge dx_1 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dx_1 \wedge dx_2. \end{aligned} \tag{6}$$

Por tanto, la 2-forma $d\omega^1$ tiene como funciones coeficientes las funciones coordenadas del campo rot F, donde $F = (F_1, F_2, F_3)$. Es claro que entonces la fórmula general de Stokes es exactamente la fórmula de Stokes de la Sec. 4 si B y ∂B son un pedazo de superficie S y su frontera orientada ∂S, respectivamente.

La fórmula de Gauss de la Sec. 5 se obtiene considerando una 2-forma

$$\omega^2 = F_1\,dx_2 \wedge dx_3 + F_2\,dx_3 \wedge dx_1 + F_3\,dx_1 \wedge dx_2.$$

Un cálculo corto muestra que

$$d\omega^2 = \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}\right) dx_1 \wedge dx_2 \wedge dx_3. \tag{7}$$

Esta 3-forma tiene como coeficiente la divergencia del campo $F = (F_1, F_2, F_3)$ y sustituyendo en la fórmula general de Stokes se obtiene la fórmula de Gauss de la Sec. 5, excepto que, tal como en la fórmula de Green, la integral de volumen de la fórmula de Gauss no es idéntica a la integral de una 3-forma. Véase Ejercicio 10.

La correspondencia entre un campo de vectores $F = (F_1, F_2, F_3)$ en $\mathfrak{R}^3$ y una forma diferencial con coeficientes F_1, F_2, F_3 ha sido descrita en las Ecs. (6) y (7) y puede ser resumida así:

si $\omega^2 \leftrightarrow F$, entonces $d\omega^2 \leftrightarrow \operatorname{div} F$

si $\omega^1 \leftrightarrow F$, entonces $d\omega^1 \leftrightarrow \operatorname{rot} F$

Finalmente, si $\omega^0 \leftrightarrow f$, entonces $d\omega^0 \leftrightarrow \operatorname{grad} f$, donde f es real.

EJERCICIOS

1. Calcule $d\omega$, donde ω es
   (a) $(x_1^2 + x_2^2)\,dx_1 - x_2\,dx_2$.
   (b) sen $x_1\,dx_3$.
   (c) $x_1\,dx_1 \wedge dx_2 + x_2\,dx_2 \wedge dx_3$.
   (d) $yz\,dx + zx\,dy + xy\,dz$.
   (e) $x_1\,dx_1 \wedge dx_2 \wedge dx_3$.
2. (a) Si $\mathfrak{R}^2 \xrightarrow{f_1} \mathfrak{R}$ y $\mathfrak{R}^2 \xrightarrow{f_2} \mathfrak{R}$ son funciones diferenciables con igual dominio, demuestre que
   $$df_1 \wedge df_2 = \frac{\partial(f_1, f_2)}{\partial(x_1, x_2)}\,dx_1 \wedge dx_2.$$
   (b) Generalice la parte (a) al caso de tres funciones de $\mathfrak{R}^3$ en $\mathfrak{R}$.

3. Demuestre que si ω^p y ω^q son una p- y una q-forma respectivamente, entonces

$$d(\omega^p \wedge \omega^q) = d\omega^p \wedge \omega^q + (-1)^p \omega^p \wedge (d\omega^q),$$

(a) si ω^p y ω^q están definidas en $\mathfrak{R}^2$,
(b) si ω^p y ω^q están definidas en $\mathfrak{R}^n$.

4. Si ω^0 es una 0-forma en $\mathfrak{R}^n$ con coeficientes dos veces continuamente diferenciables, demuestre que $d(d\omega^0) = 0$.

5. (a) Sea ω^1 una 1-forma en $\mathfrak{R}^3$ con coeficientes dos veces continuamente diferenciables. Demuestre que $d(d\omega^1) = 0$.
(b) Use la correspondencia entre un campo vectorial $F = (F_1, F_2, F_3)$ y la 1-forma con coeficientes F_1, F_2, F_3 para demostrar que el resultado de la parte (a) es equivalente a div (rot F) = 0.
(c) Demuestre que $d(d\omega^0)$ es equivalente a rot (grad f) = 0.

6. (a) Encuentre un ejemplo de una 1-forma ω^1 tal que no existe una 0-forma ω^0 tal que $d\omega^0 = \omega^1$.
(b) Demuestre en forma análoga que si ω^2 es una 2-forma, entonces no existe necesariamente una 1-forma ω^1 tal que $d\omega^1 = \omega^2$.
(c) Interprete las partes (a) y (b) en términos del gradiente y el rotor.

7. Demuestre que si ω^p es una p-forma con coeficientes dos veces diferenciables, entonces $d(d\omega^p) = 0$.

8. (a) Sea $F = (F_1, \ldots, F_n)$ un campo vectorial en $\mathfrak{R}^n$ y considere la 1-forma $\omega^1 = F_1\, dx_1 + \ldots + F_n\, dx_n$. Demuestre que la condición $d\omega^1 = 0$ es equivalente a la condición de simetría de la matriz jacobiana F'.
(b) Demuestre que si ω^1 es una 1-forma con coeficientes continuamente diferenciables en un rectángulo R en $\mathfrak{R}^n$, entonces $d\omega^1 = 0$ en R implica que existe una función $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}$ tal que $df = \omega^1$. [*Indicación*: Use el Teorema 2.6.]

9. En la fórmula general de Stokes (Ec. 8.1), sea B un disco en $\mathfrak{R}^2$ parametrizado por $(x_1, x_2) = (u \cos v,\ u \operatorname{sen} v)$, $0 \leq u \leq 1$, $0 \leq v \leq 4\pi$. Nótese que el disco está recubierto por esta parametrización. ¿Cuál es la parametrización correcta para ∂B tal que 8.1 sea válida? Considere que ω^p es una 1-forma con coeficientes continuamente diferenciables.

10. En la fórmula general de Stokes (Ec. 8.1), sea ω^p una 2-forma con coeficientes continuamente diferenciables en la bola $|\mathbf{x}| \leq 1$ en $\mathfrak{R}^3$. Demuestre que es posible parametrizar la bola B de modo tal que la parametrización correspondiente de ∂B tenga una normal que apunte hacia adentro. Explique la aparente contradicción con el requisito de que ∂B tenga una normal que apunte hacia afuera que aparece en el teorema de Gauss de la Sec. 5.

Apéndice

1. Introducción

Los teoremas y técnicas del cálculo dependen de las propiedades algebraicas y topológicas de $\mathfrak{R}^n$ y de las funciones de $\mathfrak{R}^n$ en $\mathfrak{R}^m$. Las propiedades algebraicas fueron consideradas en el Cap. 1 y las propiedades topológicas de límites y continuidad en la Sec. 7 del Cap. 3. Sin embargo, en algunas partes del desarrollo, principalmente en teoremas de existencia para integrales, necesitamos más resultados acerca de la continuidad. En esta sección consideraremos estos resultados sin dar demostraciones. Por ejemplo, consideremos el siguiente teorema.

1.1 Teorema

Sea $\mathfrak{R}^n \xrightarrow{f} \mathfrak{R}$ continua en un conjunto cerrado y acotado K en $\mathfrak{R}^n$. Entonces f alcanza su máximo y mínimo en K.

Este teorema se usa para garantizar la existencia de un máximo o un mínimo para una función real. Las técnicas habituales del cálculo son útiles para encontrar máximos y mínimos, pero no resuelven el problema de su existencia. El siguiente teorema es fundamental para las consideraciones que siguen.

1.2 Teorema

Sea K un subconjunto cerrado y acotado de $\mathfrak{R}$. Entonces K contiene un elemento más grande que todos, llamado el **supremo** de K y un elemento más pequeño que todos, llamado el **ínfimo** de K.

Para demostrar el Teorema 1.1 se puede usar el Teorema 1.2 o el resultado siguiente, que también se usa para demostrar la existencia de la integral de una función continua.

1.3 Teorema

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ continua en un subconjunto cerrado y acotado K de $\mathcal{R}^n$. Entonces f es uniformemente continua en K; es decir, dado $\epsilon > 0$ existe $\delta > 0$ tal que $|f(\mathbf{x}) - f(\mathbf{y})| < \epsilon$ cuando $\mathbf{x}$ e $\mathbf{y}$ están en K y $|\mathbf{x} - \mathbf{y}| < \delta$.

Las demostraciones de estos tres teoremas, como también la del Teorema 1.4, pueden ser encontradas en cualquier texto sobre la teoría de variables. Uno de los mejores es *Principles of Mathematical Analysis*, 2nd. Ed., Walter Rudin, McGraw-Hill, 1964.

A continuación describimos un criterio intrínseco para la convergencia de una sucesión $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ de vectores en $\mathcal{R}^n$. Decimos que la sucesión es una sucesión de Cauchy si, dado $\epsilon > 0$ existe un número N tal que

$$|\mathbf{x}_k - \mathbf{x}_l| < \epsilon$$

cuando $k, l > N$. A pesar de que es bastante fácil demostrar directamente que toda sucesión convergente es una sucesión de Cauchy, la afirmación conversa depende del Teorema 1.2.

1.4 Teorema

Si $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ es una sucesión de Cauchy de vectores en $\mathcal{R}^n$, entonces la sucesión tiene un límite $\mathbf{x}$ en $\mathcal{R}^n$, de modo que

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}.$$

El Teorema 1.4 se necesita para demostrar el Teorema 5.1, que justifica el método modificado de Newton para resolver ecuaciones vectoriales.

Las definiciones de punto límite, límite, continuidad, punto interior, conjunto abierto, frontera, conjunto cerrado y diferenciabilidad han sido dadas directa o indirectamente en término de la norma euclidiana en $\mathcal{R}^n$. Usando esta norma, la distancia entre dos puntos $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$ es el número

$$|\mathbf{x} - \mathbf{y}| = [(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2]^{1/2}.$$

Estas mismas definiciones pueden hacerse con respecto de cualquier norma. Una norma en $\mathcal{U}$ es una función real $\| \ \|$ con dominio $\mathcal{U}$ y con las tres propiedades.

1.5 Positividad: $\|\mathbf{x}\| > 0$, salvo que $\|0\| = 0$.

1.6 Homogeneidad: $\|a\mathbf{x}\| = |a| \, \|\mathbf{x}\|$.

1.7 Desigualdad triangular: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Un propósito de esta sección es demostrar que en un espacio vectorial de dimensión finita, las definiciones mencionadas son independientes de la elección de norma. Es decir, si $\mathbf{x}_0$ es un punto límite de un conjunto S con respecto de una norma, entonces también lo es con respecto de cualquier otra norma y resultados análogos para los otros conceptos. En particular, se deduce que estos conceptos no dependen de un producto interno euclidiano. Un ejemplo de una norma en $\mathcal{R}^n$ distinta de la norma euclidiana es la llamada **norma caja** o **norma del máximo** definida para todo $\mathbf{x} = (x_1, \ldots, x_n)$ por

$$\|\mathbf{x}\| = \max \{|x_1|, \ldots, |x_n|\}.$$

Si $\mathbf{x}_0 = (2, 1)$, entonces el conjunto de todos los $\mathbf{x}$ en $\mathcal{R}^2$ tales que $\|\mathbf{x} - \mathbf{x}_0\| < 0{,}5$ es el paralelogramo mostrado en la Fig. 1, donde hemos dibujado ejes no perpendiculares y con unidades de largo distintas para recalcar que ésta no es la norma euclidiana.

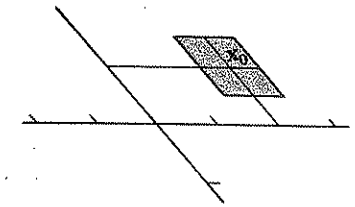


Figura 1

Dadas dos normas $\| \ \|_1$ y $\| \ \|_2$ en el espacio vectorial $\mathcal{U}$, decimos que $\| \ \|_1$ es **equivalente** a $\| \ \|_2$ si existen números reales positivos k y K tales que, para todo $\mathbf{x}$ en $\mathcal{U}$,

$$k \, \|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq K \, \|\mathbf{x}\|_1. \tag{1}$$

Es fácil comprobar que ésta es una relación de equivalencia, es decir, que satisface las condiciones de

Reflexividad: Toda norma es equivalente a sí misma.

Simetría: $\| \ \|_1$ es equivalente a $\| \ \|_2$ si, y solo si, $\| \ \|_2$ es equivalente a $\| \ \|_1$.

Transitividad: Si $\| \ \|_1$ es equivalente a $\| \ \|_2$ y equivalente a $\| \ \|_3$, entonces $\| \ \|_1$ es equivalente a $\| \ \|_3$.

Más importante es el hecho de que normas equivalentes conducen a las mismas definiciones de punto límite, límite, punto interior y diferenciabilidad. Continuidad, conjuntos abiertos, frontera y conjunto cerrado son conceptos que se definen en términos de los anteriores y, por tanto, no se alteran al cambiar la norma por una equivalente.

Para verificar lo afirmado, sean $\| \ \|_1$ y $\| \ \|_2$ dos normas equivalentes y supongamos que $\mathbf{x}_0$ es un punto límite de S con respecto de $\| \ \|_1$. Entonces

para todo $\epsilon_1 > 0$ existe un punto $\mathbf{x}$ en S tal que $0 < \|\mathbf{x} - \mathbf{x}_0\|_1 < \epsilon_1$. Por tanto, si $\epsilon_2 > 0$ es arbitrario, tomando $\epsilon_1 = \epsilon_2/K$ y usando (1), se tiene

$$0 < \|\mathbf{x} - \mathbf{x}_0\|_2 \leq K\|\mathbf{x} - \mathbf{x}_0\|_1 < K\epsilon_1 = \epsilon_2.$$

Por tanto, $\mathbf{x}_0$ es un punto límite de S con respecto de $\| \quad \|_2$. Supongamos ahora que, con respecto de $\| \quad \|_1$, tenemos que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0.$$

Entonces, tal como lo demostráramos, $\mathbf{x}_0$ es un punto límite del dominio de f con respecto de ambas normas. Para todo $\epsilon_1 > 0$ existe $\delta_1 > 0$ tal que si $\mathbf{x}$ está en el dominio de f y $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$, entonces $\|f(\mathbf{x}) - \mathbf{y}_0\|_1 < \epsilon_1$. Sea $\epsilon_2 > 0$, entonces si $\epsilon_1 = \epsilon_2/K$, podemos elegir $\delta_2 = \delta_1 k$, de modo que si $0 < \|\mathbf{x} - \mathbf{x}_0\|_2 < \delta_2$, por la desigualdad (1), tenemos que

$$0 < \|\mathbf{x} - \mathbf{x}_0\|_1 \leq \frac{1}{k}\|\mathbf{x} - \mathbf{x}_0\|_2 < \frac{\delta_2}{k} = \delta_1.$$

Por tanto,

$$\|f(\mathbf{x}) - \mathbf{y}_0\|_2 \leq K\|f(\mathbf{x}) - \mathbf{y}_0\|_1 < K\epsilon_1 = \epsilon_2,$$

y concluimos que $\lim f(\mathbf{x}) = \mathbf{y}_0$ con respecto de $\| \quad \|_2$. Los argumentos para las definiciones de punto interior y diferenciabilidad son análogos y omitimos los detalles.

La **ϵ-bola de centro $\mathbf{x}_0$** es un espacio vectorial $\mathcal{U}$ con respecto de una norma dada es el conjunto de todos los $\mathbf{x}$ en $\mathcal{U}$ tales que $\|\mathbf{x} - \mathbf{x}_0\| < \epsilon$. En general, una ϵ-bola no tiene el aspecto de una bola. Por ejemplo, con respecto de la norma caja en $\mathcal{R}^2$, toda ϵ-bola es un paralelogramo (véase Fig. 1). De la desigualdad (1) sigue que dos normas son equivalentes si, y solo si, toda ϵ-bola de centro $\mathbf{x}_0$ con respecto de una norma, está contenida en una δ-bola de centro $\mathbf{x}_0$ con respecto de la otra norma, y viceversa (véase Fig. 2). El teorema principal es el siguiente.

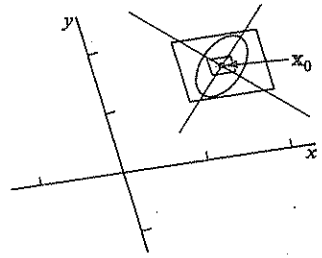


Figura 2

1.8 Teorema

Todo par de normas en un espacio vectorial de dimensión finita $\mathcal{U}$ son equivalentes.

Demostración. Sea $\| \quad \|$ una norma arbitraria en $\mathcal{U}$. Eligiendo una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ en $\mathcal{U}$, definimos una norma euclidiana en $\mathcal{U}$ por

$$|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_n^2},$$

para $\mathbf{x} = x_1\mathbf{x}_1 + \ldots + x_n\mathbf{x}_n$. Demostraremos que $\| \quad \|$ es equivalente a $| \quad |$, es decir, que existen números reales positivos k y K tales que $k|\mathbf{x}| \leq \|\mathbf{x}\| \leq K|\mathbf{x}|$ para todo $\mathbf{x}$ en $\mathcal{U}$. Por la propiedad de transitividad de la relación de equivalencia entre normas, dos normas cualesquiera en $\mathcal{U}$ son equivalentes.
Para todo $\mathbf{x} = x_1\mathbf{x}_1 + \ldots + x_n\mathbf{x}_n$, tenemos que

$$\|\mathbf{x}\| \leq \sum_{i=1}^{n} |x_i|\,\|\mathbf{x}_i\| \leq \left(\sum_{i=1}^{n} \|\mathbf{x}_i\|\right) \max_i \{|x_i|\}$$

$$\leq \left(\sum_{i=1}^{n} \|\mathbf{x}_i\|\right)\sqrt{\sum_{i=1}^{n} x_i^2} = K|\mathbf{x}|,$$

donde $K = \sum_{i=1}^{n} \|\mathbf{x}_i\| > 0$. Probamos ahora que k existe. Afirmamos que la función real $\| \quad \|$ es continua con respecto de la norma $| \quad |$ y como función de $\mathbf{x}$. En efecto, si $\epsilon > 0$ es dado, eligiendo $\delta = \epsilon/K$ se tiene que si $|\mathbf{x} - \mathbf{x}_0| < \delta$, entonces

$$|\,\|\mathbf{x}\| - \|\mathbf{x}_0\|\,| \leq \|\mathbf{x} - \mathbf{x}_0\| \leq K|\mathbf{x} - \mathbf{x}_0| < \epsilon.$$

Sea k el valor mínimo de la función $\| \quad \|$ restringida a la esfera euclidiana unitaria $|\mathbf{x}| = 1$. Entonces, para todo $\mathbf{x} \neq 0$, se tiene que $\|\mathbf{x}/|\mathbf{x}|\,\| \geq k$ y, por tanto,

$$\|\mathbf{x}\| \geq k|\mathbf{x}|, \qquad \text{para todo } \mathbf{x} \text{ en } \mathcal{U}.$$

Esto completa la demostración.

2. La regla de la cadena

En el Cap. 4 presentamos dos versiones diferentes de la regla de la cadena, una en la sección sobre el gradiente y la otra en la Sec. 3, dedicada totalmente a la regla de la cadena. Ambos teoremas tienen como hipótesis que ciertas funciones son continuamente diferenciables. Además, en el Teorema 3.1 supusimos que la función $g \circ f$ está definida en un conjunto abierto. No es necesario hacer estas suposiciones para que la regla de la cadena sea diferenciable. Por supuesto, que si queremos deducir que $g \circ f$ es continuamente diferenciable tenemos que suponer que f y g tienen matrices derivadas continuas. La virtud del teorema que damos a continuación es que él contiene un mínimo de suposiciones y su demostración es del mismo tipo de la que se daría en el caso de funciones reales de una variable.

2.1 Teorema. Regla de la cadena

Si $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ es diferenciable en $\mathbf{x}_0$ y $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^p$ es diferenciable en $f(\mathbf{x}_0)$, entonces $g \circ f$ es diferenciable en $\mathbf{x}_0$ y

$$(g \circ f)'(\mathbf{x}_0) = g'(f(\mathbf{x}_0))f'(\mathbf{x}_0).$$

Demostración. Lo primero que hay que demostrar es que $\mathbf{x}_0$ es un punto interior del dominio de $g \circ f$. Como g es diferenciable en $f(\mathbf{x}_0)$, el punto $\mathbf{y}_0 = f(\mathbf{x}_0)$ es, por definición, un punto interior del dominio de g. Por tanto, existe un número real positivo δ' tal que un punto $\mathbf{y}$ está en el dominio de g cuando $|\mathbf{y} - \mathbf{y}_0| < \delta'$. Como f es diferenciable en $\mathbf{x}_0$, también es continua en ese punto. Además, por definición, $\mathbf{x}_0$ es un punto interior del dominio de f. Por tanto, existe un número positivo δ tal que si $|\mathbf{x} - \mathbf{x}_0| < \delta$, entonces $\mathbf{x}$ está en el dominio de f y

$$|f(\mathbf{x}) - \mathbf{y}_0| = |f(\mathbf{x}) - f(\mathbf{x}_0)| < \delta'.$$

Pero δ' ha sido elegido de modo que si esta última desigualdad es válida, entonces $f(\mathbf{x})$ está en el dominio de g. Por tanto, todo punto $\mathbf{x}$ en $\mathcal{R}^n$ tal que $|\mathbf{x} - \mathbf{x}_0| < \delta$ está en el dominio de la función compuesta $g \circ f$, por tanto, el vector $\mathbf{x}_0$ es un punto interior de ese dominio.

Falta demostrar que la función lineal compuesta cuya matriz es $g'(\mathbf{y}_0)f'(\mathbf{x}_0)$ satisface las condiciones de la definición de la derivada de $g \circ f$ en $\mathbf{x}_0$. Es decir, debemos demostrar que si

$$g \circ f(\mathbf{x}) - g \circ f(\mathbf{x}_0) - g'(\mathbf{y}_0)f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = |\mathbf{x} - \mathbf{x}_0| Z(\mathbf{x} - \mathbf{x}_0), \tag{1}$$

entonces

$$\lim_{\mathbf{x} \to \mathbf{x}_0} Z(\mathbf{x} - \mathbf{x}_0) = 0.$$

Como f y g son diferenciables en $\mathbf{x}_0$ e $\mathbf{y}_0$, respectivamente, existen funciones Z_1 y Z_2 tales que

$$f(\mathbf{x}) - f(\mathbf{x}_0) = f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + |\mathbf{x} - \mathbf{x}_0| Z_1(\mathbf{x} - \mathbf{x}_0),$$
$$g(\mathbf{y}) - g(\mathbf{y}_0) = g'(\mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + |\mathbf{y} - \mathbf{y}_0| Z_2(\mathbf{y} - \mathbf{y}_0),$$

y

$$\lim_{\mathbf{x} \to \mathbf{x}_0} Z_1(\mathbf{x} - \mathbf{x}_0) = \lim_{\mathbf{y} \to \mathbf{y}_0} Z_2(\mathbf{y} - \mathbf{y}_0) = 0.$$

Sustituyendo la ecuación para f en la ecuación para g, obtenemos

$$g \circ f(\mathbf{x}) - g \circ f(\mathbf{x}_0) = g'(\mathbf{y}_0)\big(f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + |\mathbf{x} - \mathbf{x}_0| Z_1(\mathbf{x} - \mathbf{x}_0)\big)$$
$$+ |f(\mathbf{x}) - f(\mathbf{x}_0)| Z_2(f(\mathbf{x}) - f(\mathbf{x}_0)).$$

De aquí, por la linealidad de la multiplicación por la matriz $g'(\mathbf{y}_0)$, se tiene que

$$g \circ f(\mathbf{x}) - g \circ f(\mathbf{x}_0) - g'(\mathbf{y}_0)f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$
$$= |\mathbf{x} - \mathbf{x}_0| g'(\mathbf{y}_0)[Z_1(\mathbf{x} - \mathbf{x}_0)] + |f(\mathbf{x}) - f(\mathbf{x}_0)| Z_2(f(\mathbf{x}) - f(\mathbf{x}_0)). \tag{2}$$

Notamos que, como f es diferenciable en $\mathbf{x}_0$,

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = |f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + |\mathbf{x} - \mathbf{x}_0| Z_1(\mathbf{x} - \mathbf{x}_0)|$$
$$\leq k |\mathbf{x} - \mathbf{x}_0| + |Z_1(\mathbf{x} - \mathbf{x}_0)| |\mathbf{x} - \mathbf{x}_0|,$$

donde en el último paso hemos usado la desigualdad triangular y la constante k es proporcionada por el Teorema 7.3 del Cap. 3. Esta desigualdad nos permite estimar el término derecho de la Ec. (2), mostrando que su norma es menor o igual que $|\mathbf{x} - \mathbf{x}_0|$ multiplicado por una función que tiende a cero cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$. En efecto, tenemos la cota superior

$$|\mathbf{x} - \mathbf{x}_0| \{|g'(\mathbf{y}_0)[Z_1(\mathbf{x} - \mathbf{x}_0)]| + [k + |Z_1(\mathbf{x} - \mathbf{x}_0)|] |Z_2(f(\mathbf{x}) - f(\mathbf{x}_0))|\}$$

Pero cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$, $f(\mathbf{x})$ tiende a $f(\mathbf{x}_0)$, de modo que $Z_1(\mathbf{x} - \mathbf{x}_0)$ y $Z_2(f(\mathbf{x}) - f(\mathbf{x}_0))$ tienden a cero. Esto demuestra que $Z(\mathbf{x} - \mathbf{x}_0)$, definido por la Ec. (1), también tiende a cero, con lo cual queda terminada la demostración.

3. La fórmula del largo de arco

Para comprender la importancia de la conexión entre el largo de arco y la fórmula $\int_a^b |f'(t)|\, dt$ usada para calcularlo, es necesario saber que esta fórmula no siempre es válida. La razón para esto es que existe una curva continua γ en $\mathcal{R}^2$ de largo 2 tal que si se aplica la fórmula el resultado es 1. La construcción de este ejemplo es bastante complicada y la demostración de que la integral es 1 es no trivial. Como curva γ podemos tomar el gráfico de la llamada función de Cantor (véase R. P. Boas, *A Primer of Real Functions*, John Wiley and Sons, 1960, pág. 131 o B. R. Gelbaum y J. M. H. Olmstead, *Contraejemplos en Análisis*, Holden-Day, Inc., 1964, pág. 97). Una vez que se conoce bien la función de Cantor, es bastante fácil demostrar que su gráfico tiene largo 2 si se define el largo como el supremo de los largos de los polígonos inscritos.

Para una curva lisa por pedazos γ, dada por una función $\mathcal{R} \xrightarrow{g} \mathcal{R}^n$ continuamente diferenciable por pedazos para $a \leq t \leq b$, tenemos el teorema siguiente enunciado en el Cap. 3. (El gráfico de la función de Cantor no es una curva lisa por pedazos.)

Teorema

Sea γ una curva lisa por pedazos descrita de la manera recientemente indicada. Entonces el largo $l(\gamma)$ de γ es finito y

$$l(\gamma) = \int_a^b |g'(t)|\, dt.$$

Demostración. Primero demostraremos que $l(\gamma) \leq \int_a^b |g'(t)|\, dt$ y, entonces, como $|g'|$ es continua en un número finito de intervalos cerrados, ella es

acotada y la integral es finita. La desigualdad quedará demostrada si probamos que

$$\sum_{k=1}^{K}|g(t_k) - g(t_{k-1})| \leq \int_a^b |g'(t)|\, dt, \tag{1}$$

cuando $a = t_0 < t_1 < \ldots < t_K = b$ es una partición *arbitraria* P de $[a, b]$. Por la desigualdad triangular de la norma, la suma de la izquierda se hace mayor si agregamos a la partición todos los puntos extremos de los intervalos en los cuales g' es continua. Por tanto, suponemos que esto se ha hecho. Entonces, en cada intervalo $[t_{k-1}, t_k]$ tenemos que

$$\begin{aligned}|g(t_k) - g(t_{k-1})| &= \left|\int_{t_{k-1}}^{t_k} g'(t)\, dt\right| \\ &\leq \int_{t_{k-1}}^{t_k} |g'(t)|\, dt.\end{aligned}$$

La igualdad es válida por el Ejercicio 16 de la Sec. 2 del Cap. 3 y la desigualdad por el Ejercicio 17 de la misma sección. Sumando para $k = 1, \ldots, K$, se tiene (1).

Para demostrar la desigualdad inversa $\int_a^b |g'(t)|\, dt \leq l(\gamma)$, demostraremos que para todo $\eta > 0$ podemos encontrar una partición P de $[a, b]$ tal que

$$\int_a^b |g'(t)|\, dt - \eta \leq \sum_{k=1}^{K}|g(t_k) - g(t_{k-1})|. \tag{2}$$

Esto mostrará que ningún número menor que la integral es una cota superior para las sumas de la derecha. Empezamos con la partición P_0 formada por los puntos extremos de los intervalos cerrados en los cuales g' es continua. Por el Teorema 1.3 de este Apéndice, g' es uniformemente continua en cada uno de los intervalos. Esto significa que dado $\epsilon > 0$ existe $\delta > 0$ tal que $|g'(t) - g'(u)| < \epsilon$ si $|t - u| < \delta$ y $t_{k-1} \leq t, u \leq t_k$. Como hay un número finito de intervalos $[t_{k-1}, t_k]$, podemos elegir un $\delta > 0$ único que sirve en todos los intervalos. Ahora tomamos una partición P lo suficientemente fina como para que máx $(t_k - t_{k-1}) < \delta$ y que todavía contenga los puntos de P_0. En cada intervalo de la nueva partición tenemos que $|g'(t)| \leq |g'(t_k)| + \epsilon$, ya que g' es uniformemente continua. Por tanto,

$$\begin{aligned}\int_{t_{k-1}}^{t} |g'(t)|\, dt &\leq [|g'(t_k)| + \epsilon](t_k - t_{k-1}) \\ &= |g'(t_k)|\,(t_k - t_{k-1}) + \epsilon(t_k - t_{k-1}).\end{aligned} \tag{3}$$

Pero también tenemos que

$$\begin{aligned}|g'(t_k)|\,(t_k - t_{k-1}) &= \left|\int_{t_{k-1}}^{t_k} \big(g'(t) + g'(t_k) - g'(t)\big)\, dt\right| \\ &\leq \left|\int_{t_{k-1}}^{t_k} g'(t)\, dt\right| + \left|\int_{t_{k-1}}^{t_k} \big(g'(t_k) - g'(t)\big)\, dt\right|\end{aligned}$$

$$\leq |g(t_k) - g(t_{k-1})| + \int_{t_{k-1}}^{t_k} |g'(t_k) - g'(t)|\, dt,$$

donde en el último paso hemos usado nuevamente los resultados de los Ejercicios 16 y 17 de la Sec. 2 del Cap. 3. Usando la continuidad uniforme de g' y la desigualdad anterior, se tiene que

$$|g'(t_k)|\,(t_k - t_{k-1}) \leq |g(t_k) - g(t_{k-1})| + \epsilon(t_k - t_{k-1}).$$

Usando esta desigualdad en la Ec. (3), se tiene que

$$\int_{t_{k-1}}^{t_k} |g'(t)|\, dt \leq |g(t_k) - g(t_{k-1})| + 2\epsilon(t_k - t_{k-1}).$$

Sumando en k, se tiene que

$$\int_a^b |g'(t)|\, dt \leq \sum_{k=1}^{K}|g(t_k) - g(t_{k-1})| + 2\epsilon(b - a).$$

Si elegimos ϵ, de modo que $2\epsilon(b - a) < \eta$, la partición P satisface la desigualdad (2).

4. Convergencia de las series de Fourier

El Teorema 5.3 del Cap. 5 afirma que las series de Fourier de una función lisa por pedazos f converge puntualmente al promedio de los límites izquierdo y derecho de f en cada punto. La hipótesis de que f es lisa por pedazos significa que el intervalo $[-\pi, \pi]$ puede ser dividido en un número finito de subintervalos en cada uno de los cuales f y f' pueden ser extendidas en forma continua. Exigimos que en los extremos de cada intervalo $[x_k, x_{k+1}]$ f sea continua si se le asignan los valores $f(x_k+) = \lim_{u \to 0^+} f(x_k + u)$ y $f(x_{k+1}-) = \lim_{u \to 0^-} f(x_{k+1} + u)$. Análogamente, exigimos que f' sea continua en el intervalo cerrado si se le asignan los valores

$$f'(x_k) = \lim_{u \to 0+} \frac{f(x_k + u) - f(x_k +)}{u},$$

$$f'(x_{k+1}) = \lim_{u \to 0+} \frac{f(x_{k+1} + u) - f(x_{k+1} -)}{u}.$$

El teorema de la convergencia se puede expresar en la forma.

4.1 Teorema

Sea f lisa por pedazos en $[-\pi, \pi]$. Entonces la serie de Fourier de f converge en cada punto x del intervalo a $(\frac{1}{2})[f(x-) + f(x+)]$. En particular, si f es continua en x, entonces la serie converge a $f(x)$.

Demostración. Debemos demostrar que si a_k y b_k son los coeficientes de Fourier de f, entonces

$$\lim_{N\to\infty} s_N(x) = \lim_{N\to\infty}\left[\frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos kx + b_k \operatorname{sen} kx\right]$$

$$= \tfrac{1}{2}[f(x-) + f(x+)], \quad \text{para todo } x \text{ en } [-\pi, \pi].$$

Remplazando a_k y b_k por sus definiciones, tenemos que

$$s_N(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt$$

$$+ \frac{1}{\pi}\sum_{k=1}^{N}\left[\cos kx \int_{-\pi}^{\pi} f(t)\cos kt\,dt + \operatorname{sen} kx \int_{-\pi}^{\pi} f(t)\operatorname{sen} kt\,dt\right]$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\tfrac{1}{2} + \sum_{k=1}^{N}\cos k(t-x)\right] dt.$$

Por las identidades trigonométricas del Ejercicio 5 de la Sec. 5 del Cap. 5, se tiene que

$$\tfrac{1}{2} + \sum_{k=1}^{N}\cos ku = \frac{\operatorname{sen}(N+\frac{1}{2})u}{2\operatorname{sen}(\frac{1}{2})u}. \tag{1}$$

Por tanto,

$$s_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\operatorname{sen}(N+\frac{1}{2})(t-x)}{2\operatorname{sen}(\frac{1}{2})(t-x)}\,dt.$$

Extendemos f fuera del intervalo $[-\pi, \pi]$, de modo que ella tenga periodo 2π y hacemos el cambio de variables $t = x + u$. Entonces el nuevo intervalo de integración es $[-\pi - x, \pi - x]$. Pero como el integrando tiene periodo 2π, el valor de la integral no cambia si pasamos al intervalo $[-\pi, \pi]$. Por tanto, tenemos que

$$s_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+u)\frac{\operatorname{sen}(N+\frac{1}{2})u}{2\operatorname{sen}(\frac{1}{2})u}\,du.$$

Demostraremos que

$$\lim_{N\to\infty} s_N^{+}(x) = (\tfrac{1}{2})f(x+), \tag{2}$$

donde

$$s_N^{+}(x) = \frac{1}{\pi}\int_{0}^{\pi} f(x+u)\frac{\operatorname{sen}(N+\frac{1}{2})u}{2\operatorname{sen}(\frac{1}{2})u}\,du. \tag{3}$$

Un argumento similar muestra que $\lim_{N\to\infty} s_N^{-}(x) = (\frac{1}{2})f(x-)$, donde $s_N^{-} = (1/\pi)\int_{-\pi}^{0} f(x+u)\,(\operatorname{sen}(N+\frac{1}{2})u/2\operatorname{sen}(\frac{1}{2})u)\,du$, de modo que sumando ambos términos habremos terminado la demostración.

Para demostrar la Ec. (2), notamos que de la Ec. (1) se tiene que

$$\frac{1}{\pi}\int_{0}^{\pi}\frac{\operatorname{sen}(N+\frac{1}{2})u}{2\operatorname{sen}(\frac{1}{2})u}\,du = \tfrac{1}{2}.$$

Multiplicando ambos miembros por $f(x+)$ y restando a la Ec. (3) se tiene

$$s_N^{+}(x) - \tfrac{1}{2}f(x+) = \frac{1}{\pi}\int_{0}^{\pi}\frac{\operatorname{sen}(N+\frac{1}{2})u}{2\operatorname{sen}(\frac{1}{2})u}[f(x+u) - f(x+)]\,du. \tag{4}$$

La demostración se termina demostrando que esta última integral tiende a cero cuando N tiende a infinito. Pero $g(u) = [f(x+u) - f(x+)]/\operatorname{sen}(\frac{1}{2})u$ es continua por pedazos, de modo que el resultado es consecuencia de:

4.2 Lema de Riemann

Si g es continua en el intervalo $[a, b]$, entonces

$$\lim_{k\to\infty}\int_a^b g(u)\operatorname{sen} ku\,du = \lim_{k\to\infty}\int_a^b g(u)\cos ku\,du = 0.$$

Demostración. Por la desigualdad de Bessel, demostrada en la Sec. 7 del Cap. 4, tenemos que

$$\int_a^b g^2(u)\,du \geq \sum_{k=1}^{\infty}(a_k^2 + b_k^2).$$

Por tanto, la serie de la izquierda converge. De esto concluimos que a_k y b_k tienden a cero cuando k tiende a infinito.

5. Demostración de los teoremas de la función inversa y de la función implícita

Empezamos con un teorema que no solo garantiza la existencia de una función inversa, sino que también demuestra la convergencia de una modificación de la iteración de Newton, dada en 9.2 del Cap. 3.

5.1 Teorema

Sea $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n$ continuamente diferenciable en una bola cerrada de radio r y centro en $\mathbf{x}_0$. Supongamos que $|f'(\mathbf{x}_0)|^{-1}$ existe y que para un número K, $0 < K < 1$, el valor absoluto máximo de los elementos de la matriz

$$I - [f'(\mathbf{x}_0)]^{-1}f'(\mathbf{x})$$

es menor que K/n si $|\mathbf{x}_0 - \mathbf{x}| < r$. Si

$$|[f'(\mathbf{x}_0)]^{-1}f(\mathbf{x}_0)| < (1-K)r, \tag{1}$$

entonces la ecuación

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [f'(\mathbf{x}_0)]^{-1}f(\mathbf{x}_k)$$

define una sucesión $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ que converge a un vector $\bar{\mathbf{x}}$ tal que $f(\bar{\mathbf{x}}) = 0$. Además, $\bar{\mathbf{x}}$ es la única solución de $f(\mathbf{x}) = 0$ en $B_r(\mathbf{x}_0)$.

Demostración. Sea r como en la hipótesis. Sea $g(\mathbf{x}) = \mathbf{x} - [f'(\mathbf{x}_0)]^{-1}f(\mathbf{x})$. Entonces para $\mathbf{x}$ y $\mathbf{x}'$ en $B_r(\mathbf{x}_0)$ tenemos que $g(\mathbf{x}') - g(\mathbf{x}) = (\mathbf{x}' - \mathbf{x}) - [f'(\mathbf{x}_0)]^{-1}(f(\mathbf{x}') - f(\mathbf{x}))$. Aplicando el teorema del valor medio a cada función coordenada f_j de f, tenemos que

$$f_j(\mathbf{x}') - f_j(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f_j}{\partial x_i}(\mathbf{y}_j)(x_i' - x_i),$$

donde $\mathbf{x}' = (x_1', \ldots, x_n')$ y $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y}_j$ está en el segmento que une a $\mathbf{x}$ y $\mathbf{x}'$. Denotemos por $F(\mathbf{x}, \mathbf{x}')$ a la matriz

$$\left(\frac{\partial f_j}{\partial x_i}\right)_{i,\,j=1,\ldots,n},$$

tenemos que

$$f(\mathbf{x}') - f(\mathbf{x}) = F(\mathbf{x}, \mathbf{x}')(\mathbf{x}' - \mathbf{x});$$

y, por tanto,

$$g(\mathbf{x}') - g(\mathbf{x}) - (\mathbf{x}' - \mathbf{x}) - [f'(\mathbf{x}_0)]^{-1}F(\mathbf{x}, \mathbf{x}')(\mathbf{x}' - \mathbf{x}) = (I - [f'(\mathbf{x}_0)]^{-1}F(\mathbf{x}, \mathbf{x}'))(\mathbf{x}' - \mathbf{x}). \quad (2)$$

Por hipótesis, el valor absoluto máximo de los elementos de la matriz $A = I - [f'(\mathbf{x}_0)]^{-1}F(\mathbf{x}, \mathbf{x}')$ es menor o igual que K/n para cierto $0 < K < 1$. Pero para cualquier matriz $A = (a_{ij})_{i,\,j=1,\ldots,n}$ y vector $\mathbf{y} = (y_1, \ldots, y_n)$, tenemos que

$$|A\mathbf{y}| = \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}y_j\right)^2} \leq \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^2 \sum_{j=1}^{n} y_j^2\right)},$$

por la desigualdad de Cauchy-Schwarz. Por tanto,

$$|A\mathbf{y}| \leq |\mathbf{y}|\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2} \leq |\mathbf{y}|\sqrt{n^2 \max_{i,j} a_{ij}^2} = n\,|\mathbf{y}| \max_{i,j} |a_{ij}|.$$

Aplicando este resultado a la matriz $I - [f'(\mathbf{x}_0)]^{-1}F(\mathbf{x}, \mathbf{x}')$, obtenemos a partir de (2), que

$$|g(\mathbf{x}') - g(\mathbf{x})| \leq \frac{K}{n}\, n\, |\mathbf{x}' - \mathbf{x}| = K\,|\mathbf{x}' - \mathbf{x}|. \quad (3)$$

Para $k \geq 1$ y $\mathbf{x}_k$ en el dominio de g, definimos

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k).$$

Entonces

$$|\mathbf{x}_{k+1} - \mathbf{x}_k| = |g(\mathbf{x}_k) - g(\mathbf{x}_{k-1})| \leq K\,|\mathbf{x}_k - \mathbf{x}_{k-1}| \leq K^k\,|\mathbf{x}_1 - \mathbf{x}_0|. \quad (4)$$

Aplicando repetidas veces (4) para $k \geq m \geq 0$, tenemos que

$$\begin{aligned} |\mathbf{x}_{k+1} - \mathbf{x}_m| &\leq |\mathbf{x}_{k+1} - \mathbf{x}_k| + |\mathbf{x}_k - \mathbf{x}_{k-1}| + \ldots + |\mathbf{x}_{m+1} - \mathbf{x}_m| \\ &\leq (K^k + K^{k-1} + \ldots + K^m)\,|\mathbf{x}_1 - \mathbf{x}_0| \\ &\leq K^m\left(\frac{1 - K^{k-m+1}}{1 - K}\right)|\mathbf{x}_1 - \mathbf{x}_0| \\ &\leq \frac{K^m}{1 - K}\,|\mathbf{x}_1 - \mathbf{x}_0| = \frac{K^m}{1 - K}\,|[f'(\mathbf{x}_0)]^{-1}f(\mathbf{x}_0)| \\ &\leq K^m r \end{aligned}$$

Para $m = 0$, obtenemos

$$|\mathbf{x}_{k+1} - \mathbf{x}_0| \leq r;$$

de modo que $\mathbf{x}_k$ está en $B_r(\mathbf{x}_0)$ para $k = 1, 2, 3, \ldots$, es decir, $\mathbf{x}_k$ está en el subconjunto del dominio de f para el cual la hipótesis del teorema es satisfecha. Por tanto, cada $\mathbf{x}_k$ satisface la desigualdad (4).

Volviendo a (4), vemos que

$$|\bar{\mathbf{x}}_{k+1} - \mathbf{x}_m| \leq K^m r$$

implica que $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ es una sucesión de Cauchy que converge a un vector $\bar{\mathbf{x}}$ que satisface $|\bar{\mathbf{x}} - \mathbf{x}_0| < r$. Como

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

y g es continua, tenemos que

$$\bar{\mathbf{x}} = \lim_{k\to\infty} \mathbf{x}_{k+1} = \lim_{k\to\infty} g(\mathbf{x}_k) = g(\bar{\mathbf{x}}),$$

es decir,

$$\bar{\mathbf{x}} = \bar{\mathbf{x}} - [f'(\mathbf{x}_0)]^{-1}f(\bar{\mathbf{x}}).$$

Por tanto,

$$f(\bar{\mathbf{x}}) = 0.$$

Para demostrar que la solución $\bar{\mathbf{x}}$ es única en $B_r(\mathbf{x}_0)$, notamos que la desigualdad (3), reescrita en términos de f, es

$$|(\mathbf{x}' - \mathbf{x}) - [f'(\mathbf{x}_0)]^{-1}(f(\mathbf{x}') - f(\mathbf{x}))| \leq K\,|\mathbf{x}' - \mathbf{x}|.$$

Por la desigualdad triangular se tiene que

$$|\mathbf{x}' - \mathbf{x}| - |[f'(\mathbf{x}_0)]^{-1}(f(\mathbf{x}') - f(\mathbf{x}))| \leq K\,|\mathbf{x}' - \mathbf{x}|.$$

Por tanto,

$$(1 - K)\,|\mathbf{x}' - \mathbf{x}| \leq |[f'(\mathbf{x}_0)]^{-1}(f(\mathbf{x}') - f(\mathbf{x}))|. \quad (5)$$

Esto muestra que para $\mathbf{x}$ y $\mathbf{x}'$ en $B_r(\mathbf{x}_0)$, $f(\mathbf{x}') - f(\mathbf{x}) = 0$ implica $\mathbf{x}' - \mathbf{x} = 0$, es decir, que f es inyectiva en $B_r(\mathbf{x}_0)$.

5.2 Teorema de la función inversa

Sea $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^n$ una función continuamente diferenciable tal que $f'(\mathbf{x}_0)$ tiene inversa. Entonces existe un conjunto abierto N que contiene $\mathbf{x}_0$ tal que $f(N)$ es abierto y f restringido a N tiene una inversa continuamente diferenciable f^{-1}. Además,

$$[f^{-1}]'(\mathbf{y}_0) = [f'(\mathbf{x}_0)]^{-1},$$

donde $\mathbf{y}_0 = f(\mathbf{x}_0)$.

Demostración. Parte 1. Aplicamos el teorema anterior a la función $f(\mathbf{x}) - \mathbf{y}$, donde $\mathbf{y}$ es cualquier vector tal que la desigualdad (1) es satisfecha. En otras palabras, suponemos que $\mathbf{y}$ está en el conjunto S de los vectores $\mathbf{z}$ que satisfacen

$$|[f'(\mathbf{x}_0)]^{-1}(f(\mathbf{x}_0) - \mathbf{z})| < (1 - K)r.$$

Por el teorema de aproximación, existe un vector único $\bar{\mathbf{x}}$ en $B_r(\mathbf{x}_0)$ tal que $f(\bar{\mathbf{x}}) - \mathbf{y} = 0$.

Además, la desigualdad (1) muestra que la función f^{-1} definida por $f^{-1}(\mathbf{y}) = \bar{\mathbf{x}}$ es continua puesto que

$$(1 - K)\,|f^{-1}(\mathbf{y}') - f^{-1}(\mathbf{y})| \leq |[f'(\mathbf{x}_0)]^{-1}(\mathbf{y}' - \mathbf{y})|.$$

Definimos $N = f^{-1}(S)$. Es claro que $N \subset B_r(\mathbf{x}_0)$ y es fácil comprobar que N y $f(N)$ son abiertos puesto que f y f^{-1} son continuas.

Parte 2. Para todo $\mathbf{x}$ *en* N, *la inversa de* $f'(\mathbf{x})$ *es la derivada de* f^{-1} *en* $\mathbf{y} = f(\mathbf{x})$.

En efecto, si $\mathbf{x}'$ y $\mathbf{x}$ están en N, sea $\mathbf{y}' = f(\mathbf{x}')$ e $\mathbf{y} = f(\mathbf{x})$. Como f es diferenciable en $\mathbf{x}$,

$$f(\mathbf{x}') - f(\mathbf{x}) = f'(\mathbf{x})(\mathbf{x}' - \mathbf{x}) + |\mathbf{x}' - \mathbf{x}|\,Z(\mathbf{x}' - \mathbf{x}),$$

donde $\lim_{\mathbf{x}' \to \mathbf{x}} Z(\mathbf{x}' - \mathbf{x}) = 0$. También podemos escribir la relación anterior en la forma

$$\mathbf{y}' - \mathbf{y} = f'(\mathbf{x})(f^{-1}(\mathbf{y}') - f^{-1}(\mathbf{y})) + |\mathbf{x}' - \mathbf{x}|\,Z(\mathbf{x}' - \mathbf{x}).$$

Aplicando $[f'(\mathbf{x})]^{-1}$ a ambos lados se tiene

$$f^{-1}(\mathbf{y}') - f^{-1}(\mathbf{y}) - [f'(\mathbf{x})]^{-1}(\mathbf{y}' - \mathbf{y}) = |\mathbf{x}' - \mathbf{x}|\,[f'(\mathbf{x})]^{-1}Z(\mathbf{x}' - \mathbf{x}).$$

El Lema 5.3 demuestra que $|\mathbf{x}' - \mathbf{x}|/|\mathbf{y}' - \mathbf{y}|$ se mantiene acotado cuando $\mathbf{y}'$ tiende a $\mathbf{y}$. Como $[f'(\mathbf{x})]^{-1}Z(\mathbf{x}' - \mathbf{x})$ tiende a cero cuando $\mathbf{x}'$ tiende a $\mathbf{x}$, la desigualdad $|\mathbf{y}' - \mathbf{y}| \geq M|\mathbf{x}' - \mathbf{x}|$ que resulta del Lema 5.3, muestra que el miembro de la derecha de la ecuación anterior tiende a cero cuando se lo divide por $|\mathbf{y}' - \mathbf{y}|$ y se hace tender $\mathbf{y}'$ a $\mathbf{y}$. Esto termina la demostración de la parte 2.

Parte 3. f^{-1} *es continuamente diferenciable.* La parte 2 muestra que f^{-1} es diferenciable en $f(N)$ y que

$$[f^{-1}]'(\mathbf{y}) = [f'(\mathbf{x})]^{-1}.$$

El Teorema 8.3 del Cap. 1 muestra que la inversa de una matriz A tiene elementos que son funciones continuas de los elementos de A. Como $f'(f^{-1}(\mathbf{y}))$ tiene elementos continuos, su inversa también los tiene. Esto completa la demostración del teorema.

5.3 Lema

Existe una vecindad N de $\mathbf{x}_0$ y un número $M > 0$ tal que

$$M\,|x' - x| \leq |f(x') - f(x)|$$

para todo $\mathbf{x}$ y $\mathbf{x}'$ en N.

Demostración. Por el Teorema 7.3 del Cap. 3, existe una constante k tal que

$$|[f'(x_0)]^{-1}(f(x') - f(x))| \leq k\,|f(x') - f(x)|.$$

Combinando este resultado con la desigualdad (5), se obtiene el resultado tomando $M = (1 - K)/k$.

5.4 Teorema de la función implícita

Sea $\mathcal{R}^{n+m} \xrightarrow{F} \mathcal{R}^m$ una función continuamente diferenciable. Supongamos que para algún $\mathbf{x}_0$ en $\mathcal{R}^n$ e $\mathbf{y}$ en $\mathcal{R}^m$,

1. $F(\mathbf{x}_0, \mathbf{y}_0) = 0$.
2. $F_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$ tiene inversa.

Entonces existe una función continuamente diferenciable $\mathcal{R}^n \xrightarrow{f} \mathcal{R}^m$ definida en una vecindad N de $\mathbf{x}_0$ tal que $f(\mathbf{x}_0) = \mathbf{y}_0$ y $F(\mathbf{x}, f(\mathbf{x})) = 0$, para todo $\mathbf{x}$ en N.

Demostración. La demostración consiste en reducir este teorema a una aplicación del teorema de la función inversa. Para ello, extendemos F a una función $\mathcal{R}^{n+m} \xrightarrow{H} \mathcal{R}^{n+m}$ haciendo $H(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$. Las funciones coordenadas de H están dadas en términos de las funciones coordenadas de F: $F_1, \ldots, F_m$, por

$$\begin{aligned}
H_1(\mathbf{x}, \mathbf{y}) &= x_1 \\
H_2(\mathbf{x}, \mathbf{y}) &= x_2 \\
&\cdots \\
H_n(\mathbf{x}, \mathbf{y}) &= x_n \\
H_{n+1}(\mathbf{x}, \mathbf{y}) &= F_1(x_1, \ldots, x_n, y_1, \ldots, y_m) \\
&\cdots \\
H_{n+m}(\mathbf{x}, \mathbf{y}) &= F_m(x_1, \ldots, x_n, y_1, \ldots, y_m).
\end{aligned}$$

La matriz jacobiana de H en $(\mathbf{x}_0, \mathbf{y}_0)$ es

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ 0 & 0 & & 1 & 0 & & 0 \\ \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_1}{\partial x_2} & \dots & \dfrac{\partial F_1}{\partial x_n} & \dfrac{\partial F_1}{\partial y_1} & \dots & \dfrac{\partial F_1}{\partial y_m} \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \dfrac{\partial F_m}{\partial x_1} & \dfrac{\partial F_m}{\partial x_2} & \dots & \dfrac{\partial F_m}{\partial x_n} & \dfrac{\partial F_m}{\partial y_1} & \dots & \dfrac{\partial F_m}{\partial y_m} \end{bmatrix},$$

donde todas las derivadas parciales están evaluadas en $(\mathbf{x}_0, \mathbf{y}_0)$. Por la hipótesis 2, las m columnas que aparecen a la derecha de esta matriz, son independientes. Como ellas también son independientes de las n columnas de la izquierda, todas las columnas son independientes y, por tanto, la diferencial de H en $(\mathbf{x}_0, \mathbf{y}_0)$ tiene inversa.

Es evidente que la función H es continuamente diferenciable, razón por la cual podemos aplicar en $(\mathbf{x}_0, \mathbf{y}_0)$ el teorema de la función inversa y obtener así una función H^{-1} inversa de H en un conjunto abierto N' en $\mathcal{R}^{n+m}$ que contiene a $H(\mathbf{x}_0, \mathbf{y}_0)$ y con imagen un conjunto abierto que contiene a $(\mathbf{x}_0, \mathbf{y}_0)$. Como $H(\mathbf{x}_0, \mathbf{y}_0) = (\mathbf{x}_0, F(\mathbf{x}_0, \mathbf{y}_0)) = (\mathbf{x}_0, 0)$, el conjunto N de todos los puntos $\mathbf{x}$ de $\mathcal{R}^n$ tales que $(\mathbf{x}, 0)$ está en N' es un conjunto abierto (véase el Ejercicio 22) que contiene a $\mathbf{x}_0$.

Sea G_1 la función que elige las primeras n variables de un punto de $\mathcal{R}^{n+m}$ y G_2 la función que elige las últimas m variables. Por tanto, $G_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}$ y $G_2(\mathbf{x}, \mathbf{y}) = \mathbf{y}$. Como $H(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$, la función H es la identidad en $\mathbf{x}$. Por tanto, lo mismo debe ser verdad para H^{-1}. Entonces

$$G_1 = G_1 \circ H^{-1}.$$

Definimos f por

$$f(\mathbf{x}) = G_2 H^{-1}(\mathbf{x}, 0), \quad \text{para todo } \mathbf{x} \text{ en } N.$$

Entonces

$$H^{-1}(\mathbf{x}, 0) = (G_1 H^{-1}(\mathbf{x}, 0), G_2 H^{-1}(\mathbf{x}, 0)) = (\mathbf{x}, f(\mathbf{x})).$$

Aplicando H a ambos lados, obtenemos

$$(\mathbf{x}, 0) = H(\mathbf{x}, f(\mathbf{x})) = (\mathbf{x}, f(\mathbf{x})),$$

para todo $\mathbf{x}$ en N. Las dos componentes del primer y último par deben ser iguales. Por tanto,

$$0 = F(\mathbf{x}, f(\mathbf{x})), \quad \text{para todo } \mathbf{x} \text{ en } N.$$

Por último, como f es la composición de dos funciones continuamente diferenciables, por la regla de la cadena, ella es continuamente diferenciable.

6. Demostración del teorema de Lagrange

La demostración usa el teorema de la función implícita y el hecho de que, para una función lineal L definida en $\mathcal{R}^n$, la dimensión del dominio es igual a la dimensión de la imagen sumada a la dimensión de su núcleo. Empezamos por reescribir el enunciado.

Teorema de Lagrange

Sea $\mathcal{R}^n \xrightarrow{G} \mathcal{R}^m$, $n > m$, una función continuamente diferenciable con funciones coordenadas $G_1, G_2, \dots, G_m$. Supongamos que las ecuaciones

$$\begin{aligned} G_1(x_1, \dots, x_n) &= 0 \\ G_2(x_1, \dots, x_n) &= 0 \\ &\cdot \\ &\cdot \\ &\cdot \\ G_m(x_1, \dots, x_n) &= 0 \end{aligned}$$

definen implícitamente una superficie S en $\mathcal{R}^n$ y que en un punto $\mathbf{x}_0$ de S la matriz $G'(\mathbf{x}_0)$ tiene m columnas linealmente independientes.

Si $\mathbf{x}_0$ es un punto extremo de una función diferenciable $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ restringida a S, entonces $\mathbf{x}_0$ es un punto crítico de la función

$$f + \lambda_1 G_1 + \dots + \lambda_m G_m$$

para ciertas constantes $\lambda_1, \dots, \lambda_m$.

Demostración. El teorema de la función implícita asegura la existencia de una representación paramétrica de S en una vecindad de $\mathbf{x}_0$. En efecto, supongamos que para una cierta elección de m variables, digamos $x_1, \dots, x_m$, las columnas de la matriz

$$\begin{bmatrix} \dfrac{\partial G_1}{\partial x_1} & \dfrac{\partial G_1}{\partial x_2} & \dots & \dfrac{\partial G_1}{\partial x_m} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \dfrac{\partial G_m}{\partial x_1} & \dfrac{\partial G_m}{\partial x_2} & \dots & \dfrac{\partial G_m}{\partial x_m} \end{bmatrix}_{\mathbf{x}_0} \tag{1}$$

son independientes. Entonces la matriz tiene una inversa. Escribamos $\mathbf{x}_0 = (a_1, \dots, a_n)$ y hagamos $\mathbf{u}_0 = (a_1, \dots, a_m)$, $\mathbf{v}_0 = (a_{n+1}, \dots, a_n)$. Por el teorema

de la función implícita, existe una función diferenciable $\mathcal{R}^{n-m} \xrightarrow{h} \mathcal{R}^m$ definida en una vecindad N de $\mathbf{v}_0$, tal que $h(\mathbf{v}_0) = \mathbf{u}_0$ y $G(h(\mathbf{v}), \mathbf{v}) = 0$ para todo $\mathbf{v}$ en N. La función $\mathcal{R}^{n-m} \xrightarrow{H} \mathcal{R}^n$ definida por

$$H(\mathbf{v}) = \big(h(\mathbf{v}), \mathbf{v}\big) \quad \text{para todo } \mathbf{v} \text{ en } N,$$

es una representación paramétrica de una parte de S que contiene a $\mathbf{x}_0 = H(\mathbf{v}_0)$. La superficie S tiene una tangente $\mathcal{T}$ de dimensión $n - m$ en $\mathbf{x}_0$, ya que la derivada de H en $\mathbf{v}_0$ es la matriz $(n-m)$-por-$(n-m)$

$$\begin{bmatrix} \dfrac{\partial h_1}{\partial x_{m+1}} & \dfrac{\partial h_1}{\partial x_{m+2}} & \cdots & \dfrac{\partial h_1}{\partial x_n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \dfrac{\partial h_m}{\partial x_{m+1}} & \cdots & & \dfrac{\partial h_m}{\partial x_n} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{\mathbf{v}_0},$$

donde $h_1, \ldots, h_m$ son las funciones coordenadas de H. Además, las columnas de esta matriz son independientes, ya que las columnas de ceros y unos son independientes.

Ahora compongamos H con f. Como $\mathbf{x}_0$ es un punto extremo de f en S, $\mathbf{v}$ es un punto extremo de $f \circ H$. Por tanto,

$$(f \circ H)'(\mathbf{v}_0) = f'(\mathbf{x}_0)H'(\mathbf{v}_0) = 0. \tag{2}$$

Como G es constante igual a cero en S,

$$(G \circ H)'(\mathbf{v}_0) = G'(\mathbf{x}_0)H'(\mathbf{v}_0) = 0. \tag{3}$$

Mirando a (2) y (3), vemos que $d_{\mathbf{x}_0}f$ y $d_{\mathbf{x}_0}G$ son ambas cero en la imagen de $d_{\mathbf{v}_0}H$, la cual es la tangente $\mathcal{T}$. Por tanto, la matriz

$$\begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \\ \dfrac{\partial G_1}{\partial x_1} & \cdots & \dfrac{\partial G_1}{\partial x_n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \dfrac{\partial G_m}{\partial x_1} & \cdots & \dfrac{\partial G_m}{\partial x_n} \end{bmatrix}_{\mathbf{x}_0}$$

define una función lineal $\mathcal{R}^n \xrightarrow{L} \mathcal{R}^{m+1}$ que es idénticamente cero en $\mathcal{T}$. Como la dimensión de $\mathcal{T}$ es $n - m$, tenemos que

$$n - m \leq \text{dimensión del núcleo de } L.$$

Para una función lineal L es siempre verdadero que

$$n = \text{dimensión del núcleo de } L + \text{dimensión de la imagen de } L,$$

luego,

$$n \geq n - m + \text{dimensión de la imagen de } L,$$

es decir,

$$m \geq \text{dimensión de la imagen de } L.$$

Recordemos que la imagen de L es un subespacio de $\mathcal{R}^{m+1}$. Por tanto, podemos definir una función lineal $\mathcal{R}^{m+1} \xrightarrow{\Lambda} \mathcal{R}^n$ tal que Λ es cero en la imagen de L, pero que no es idénticamente cero en $\mathcal{R}^{m+1}$. En otras palabras, existe Λ distinto de cero tal que $\Lambda \circ L = 0$. En forma matricial, $\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, de modo que

$$(\lambda_0, \lambda_1, \ldots, \lambda_m) \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \\ \dfrac{\partial G_1}{\partial x_1} & & \dfrac{\partial G_1}{\partial x_n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \dfrac{\partial G_m}{\partial x_1} & \cdots & \dfrac{\partial G_m}{\partial x_n} \end{bmatrix}_{\mathbf{x}_0} = 0. \tag{4}$$

No puede suceder que $\lambda_0 = 0$, ya que en ese caso las filas de (1) serían dependientes, lo cual contradice el hecho de que (1) tiene una inversa. Tomando $\lambda_0 = 1$ (si $\lambda_0 \neq 1$, dividimos todos los λ_i por λ_0), la condición (4) toma la forma

$$\frac{\partial f}{\partial x_j}(\mathbf{x}_0) + \lambda_1 \frac{\partial G_1}{\partial x_j}(\mathbf{x}_0) + \ldots + \lambda_m \frac{\partial G_m}{\partial x_j}(\mathbf{x}_0) = 0,$$

para $j = 1, \ldots, m$. En otras palabras $(f + \lambda_1 G_1 + \ldots + \lambda_m G_m)'(\mathbf{x}_0) = 0$. Esto termina la demostración.

7. Demostración del teorema de Taylor

El método de demostración consiste en la reducción del problema al caso de una variable en el cual se puede hacer una estimación de la fórmula integral para el resto.

7.1 Teorema de Taylor

Supongamos que $\mathcal{R}^n \xrightarrow{f} \mathcal{R}$ tiene derivadas de orden N continuas en una vecindad de $\mathbf{x}_0$. Sea $T_N(\mathbf{x} - \mathbf{x}_0)$ el desarrollo de Taylor de grado N de f alrededor de $\mathbf{x}_0$. Es decir,

$$T_N(\mathbf{x} - \mathbf{x}_0) = f(\mathbf{x}_0) + d_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0) + \ldots + \frac{1}{N!} d^N_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0).$$

Entonces

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{(f(\mathbf{x}) - T_N(\mathbf{x} - \mathbf{x}_0))}{|\mathbf{x} - \mathbf{x}_0|^N} = 0, \tag{1}$$

y T_N es el único polinomio de grado N con esta propiedad.

Demostración. Sea $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$ y definamos

$$F(t) = f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) = f(\mathbf{x}_0 + t\mathbf{y}).$$

Entonces para $k = 0, 1, \ldots, N$ podemos aplicar la regla de la cadena para obtener

$$F^{(k)}(t) = d^k_{\mathbf{x}_0 + t\mathbf{y}} f(\mathbf{y}). \tag{2}$$

Para comprobar esto, notemos que para $k = 0$, la fórmula es válida por definición. Suponiendo que ella es válida para algún $k < N$, tenemos que

$$\begin{aligned} F^{(k+1)}(t) &= \frac{d}{dt} d^k_{\mathbf{x}_0 + t\mathbf{y}} f(\mathbf{y}) \\ &= \frac{d}{dt}\left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^k_{\mathbf{x}_0 + t\mathbf{y}} f \\ &= d_{\mathbf{x}_0 + t\mathbf{y}}\left[\left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^k f\right](\mathbf{y}) \\ &= \left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^{k+1}_{\mathbf{x}_0 + t\mathbf{y}} f \\ &= d^{k+1}_{\mathbf{x}_0 + t\mathbf{y}} f(\mathbf{y}). \end{aligned}$$

lo cual termina la demostración por inducción de la Ec. (2). En particular,

$$F^{(k)}(0) = d^k_{\mathbf{x}_0} f(\mathbf{y}).$$

Del Teorema 3.1 del Cap. 5, se deduce que

$$\begin{aligned} F(1) - F(0) - \frac{1}{1!} F'(0) - \ldots - \frac{1}{N!} F^{(N)}(0) \\ = \frac{1}{(N-1)!} \int_0^1 (1-t)^{N-1} [F^{(N)}(t) - F^{(N)}(0)]\, dt. \end{aligned}$$

En términos de f, se tiene que

$$\begin{aligned} f(\mathbf{x}) - T_N(\mathbf{x} - \mathbf{x}_0) &= f(\mathbf{x}) - f(\mathbf{x}_0) - \frac{1}{1!} d_{\mathbf{x}_0} f(\mathbf{y}) - \ldots - \frac{1}{N!} d^N_{\mathbf{x}_0} f(\mathbf{y}) \\ &= \frac{1}{(N-1)!} \int_0^1 (1-t)^{N-1} [d^N_{\mathbf{x}_0 + t\mathbf{y}} f(\mathbf{y}) - d^N_{\mathbf{x}_0} f(\mathbf{y})]\, dt. \end{aligned}$$

Ahora hacemos una estimación de la diferencia

$$\begin{aligned} |f(\mathbf{x}) - T_N(\mathbf{y})| &\le \frac{1}{(N-1)!} \max_{0\le t\le 1} |d^N_{\mathbf{x}_0 + t\mathbf{y}} f(\mathbf{y}) - d^N_{\mathbf{x}_0} f(\mathbf{y})| \\ &\le \max_{0\le t\le 1} \left| \left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^N_{\mathbf{x}_0 + t\mathbf{y}} f \right. \\ &\qquad \left. - \left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^N_{\mathbf{x}_0} f \right| \\ &= \max_{0\le t\le 1} \left| \sum_{k_1 + \ldots + k_n = N} \binom{N}{k_1 \ldots k_n} y_1^{k_1} \ldots y_n^{k_n} \right. \\ &\qquad \left. \cdot \left(\frac{\partial^N f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} (\mathbf{x}_0 + t\mathbf{y}) - \frac{\partial_N f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} (\mathbf{x}_0) \right) \right|. \end{aligned}$$

Entonces, como $|y_i| \le |\mathbf{y}|$, tenemos que

$$\frac{|y_1^{k_1} \ldots y_n^{k_n}|}{|\mathbf{y}|^N} \le 1,$$

de modo que

$$\frac{|f(\mathbf{x}) - T_N(\mathbf{y})|}{|\mathbf{y}|^N} \le \tag{3}$$

$$\sum_{k_1 + \ldots + k_n = N} \binom{N}{k_1 \ldots k_n} \cdot \max_{0\le t\le 1} \left| \frac{\partial^N f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} (\mathbf{x}_0 + t\mathbf{y}) - \frac{\partial^N f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} (\mathbf{x}_0) \right|.$$

Por hipótesis, las derivadas de f de orden menor o igual que N son continuas en $\mathbf{x}_0$. Entonces, cuando $\mathbf{y}$ tiende a cero, cada término de la última suma tiende a cero, lo cual demuestra la Ec. (1). La desigualdad (3) muestra que si f es un polinomio de grado N, entonces él es igual al desarrollo de Taylor de grado N. En efecto, en tal caso todos los términos de la derecha son cero.

La demostración de que T_N es el único polinomio de grado N que satisface la Ec. (1) es como sigue. Sean T_N y T'_N dos de estos polinomios. Por (1),

$$\lim_{\mathbf{y}\to 0} \frac{T_N(\mathbf{y}) - T'_N(\mathbf{y})}{|\mathbf{y}|^N} = 0.$$

Supongamos que $T_N - T'_N$ no es idénticamente cero, y sea

$$P_k(\mathbf{y}) + R(\mathbf{y}) = T_N(\mathbf{y}) - T'_N(\mathbf{y}),$$

donde P_k es el polinomio formado por los términos del grado más pequeño que aparece en $T_N - T'_N$ (denotamos por k este grado). Entonces, existe un vector $\mathbf{y}_0$ tal que $P_k(\mathbf{y}_0) \neq 0$. Por otra parte, como $k \leq N$,

$$0 = \lim_{t\to 0} \frac{T_N(t\mathbf{y}_0) - T'_N(t\mathbf{y}_0)}{|t\mathbf{y}_0|^k}$$

$$= \lim_{t\to 0} \frac{P_k(t\mathbf{y}_0) + R(t\mathbf{y}_0)}{|t\mathbf{y}_0|^k}$$

$$= \frac{P_k(\mathbf{y}_0)}{|\mathbf{y}_0|^k} + \lim_{t\to 0} \frac{R(t\mathbf{y}_0)}{|t|^k\,|\mathbf{y}_0|^k}.$$

Sin embargo, como todos los términos de R son de grado mayor que k, el último límite es cero. Pero entonces $P_k(\mathbf{y}_0) = 0$, lo cual es una contradicción.

8. Existencia de la integral de Riemann

El Teorema 2.1 del Cap. 6 es el siguiente:

Teorema

Sea f definida y acotada en un conjunto acotado B en $\mathcal{R}^n$ y supongamos que la frontera de B está contenida en un número finito de conjuntos lisos. Si f es continua en B, salvo posiblemente en un número finito de conjuntos lisos, entonces f es integrable en B. El valor de $\int_B f\,dV$ no varía si se cambian los valores de f en un conjunto liso.

Recordemos que un **conjunto liso** en $\mathcal{R}^n$ es la imagen de un conjunto cerrado por una función continuamente diferenciable $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^n$ con $m < n$. Demostraremos primero por qué los conjuntos lisos son negligibles en el dominio de integración de f.

8.1 Teorema

Sea $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^n$ continuamente diferenciable. Entonces para todo conjunto cerrado y acotado K en el dominio de g existe una constante M tal que

$$|g(\mathbf{y}) - g(\mathbf{x})| \leq M\,|\mathbf{x} - \mathbf{y}|$$

para todo $\mathbf{x}$ e $\mathbf{y}$ en K.

Demostración. Denotemos por $K \times K$ el subconjunto de $\mathcal{R}^{2m}$ formado por todas las $2m$-plas $(\mathbf{x}, \mathbf{y})$ tales que $\mathbf{x}$ e $\mathbf{y}$ están en K. Es fácil ver que $K \times K$ es cerrado y acotado en $\mathcal{R}^{2m}$. Ahora consideramos la ecuación

$$g(\mathbf{y}) - g(\mathbf{x}) - g'(\mathbf{x})(\mathbf{x} - \mathbf{y}) = |\mathbf{x} - \mathbf{y}|\,Z(\mathbf{y} - \mathbf{x}).$$

Esta ecuación define Z, excepto en 0, de modo que definimos $Z(0) = 0$. Primero demostraremos que $Z(\mathbf{x} - \mathbf{y})$ así definido es una función continua en $K \times K$.

Para despejar $Z(\mathbf{y} - \mathbf{x})$, dividimos por $|\mathbf{x} - \mathbf{y}|$ ambos miembros de la ecuación que lo define. Entonces, como g y g' son continuas, es claro que $Z(\mathbf{x} - \mathbf{y})$ es continua con la posible excepción de $\mathbf{x} = \mathbf{y}$. Si $\mathbf{x}$ e $\mathbf{y}$ tienden a un punto $\mathbf{z}$ del dominio de g, entonces $(\mathbf{x}, \mathbf{y})$ tiende a $(\mathbf{z}, \mathbf{z})$ y queremos demostrar que $Z(\mathbf{x} - \mathbf{y})$ tiende a cero. Aplicamos el teorema del valor medio (Teorema 8.2 del Cap. 2) a las funciones coordenadas g_k de g, para $k = 1, \ldots, n$. Tenemos que, para cada k,

$$g_k(\mathbf{y}) - g_k(\mathbf{x}) = g'_k(\mathbf{x}_k)(\mathbf{x} - \mathbf{y}),$$

para $\mathbf{x}$ e $\mathbf{y}$ en una vecindad suficientemente pequeña de $\mathbf{z}$ y para algún $\mathbf{x}_k$ en el segmento que une $\mathbf{x}$ a $\mathbf{y}$. Entonces

$$\begin{aligned}
|\mathbf{x} - \mathbf{y}|\,Z(\mathbf{x} - \mathbf{y})| &= |g(\mathbf{y}) - g(\mathbf{x}) - g'(\mathbf{x})(\mathbf{y} - \mathbf{x})| \\
&\leq \max_{1\leq k\leq n} |g_k(\mathbf{y}) - g_k(\mathbf{x}) - g'_k(\mathbf{x})(\mathbf{y} - \mathbf{x})| \\
&\leq \max_{1\leq k\leq n} |g'_k(\mathbf{x}_k)(\mathbf{y} - \mathbf{x}) - g'_k(\mathbf{x})(\mathbf{y} - \mathbf{x})| \\
&= \max_{1\leq k\leq n} \left| \sum_{j=1}^{m} \left(\frac{\partial g_k}{\partial x_j}(\mathbf{x}_j) - \frac{\partial g_k}{\partial x_j}(\mathbf{x}) \right)(y_j - x_j) \right| \\
&\leq \max_{1\leq k\leq n} \left| \sum_{j=1}^{m} \frac{\partial g_k}{\partial x_j}(\mathbf{x}_k) - \frac{\partial g_k}{\partial x_j}(\mathbf{x}) \right| |\mathbf{y} - \mathbf{x}|.
\end{aligned}$$

Por tanto,

$$|Z(\mathbf{y} - \mathbf{x})| \leq \max_{1\leq k\leq n} \left| \sum_{j=1}^{m} \frac{\partial g_k}{\partial x_j}(\mathbf{x}_k) - \frac{\partial g_k}{\partial x_j}(\mathbf{x}) \right|.$$

Como por hipótesis las derivadas parciales son continuas y como cada $\mathbf{x}_k$ tiende a $\mathbf{z}$ cuando $\mathbf{x}$ e $\mathbf{y}$ tienden a $\mathbf{z}$, se tiene que $Z(\mathbf{y} - \mathbf{x})$ tiende a cero cuando $\mathbf{x}$ e $\mathbf{y}$ tiende a $\mathbf{z}$. Esto muestra que $Z(\mathbf{x} - \mathbf{y})$ es continua en $K \times K$.

Como $|Z(\mathbf{x} - \mathbf{y})|$ es continua, ella alcanza su valor máximo M', es decir, $|Z(\mathbf{y} - \mathbf{x})| \leq M'$ para todo $\mathbf{x}$ e $\mathbf{y}$ en K. Por tanto,

$$|g(\mathbf{y}) - g(\mathbf{x}) - g'(\mathbf{x})(\mathbf{y} - \mathbf{x})| \leq M'\,|\mathbf{y} - \mathbf{x}|.$$

La desigualdad $|A| - |B| \leq |A - B|$ muestra que

$$|g(\mathbf{y}) - g(\mathbf{x})| \leq M'\,|\mathbf{y} - \mathbf{x}| + |g'(\mathbf{x})(\mathbf{y} - \mathbf{x})|. \tag{1}$$

Pero

$$|g'(\mathbf{x})(\mathbf{y} - \mathbf{x})| \leq \max_{1\leq k\leq n} \sum_{j=1}^{m} \frac{\partial g_k}{\partial x_j}(\mathbf{x})\,|\mathbf{y} - \mathbf{x}|,$$

y la continuidad de las derivadas parciales en K implica la existencia de una constante M'' tal que

$$|g'(\mathbf{x})(\mathbf{y} - \mathbf{x})| \leq M''\,|\mathbf{y} - \mathbf{x}|.$$

Esta desigualdad y (1) implican que

$$|g(\mathbf{y}) - g(\mathbf{x})| \leq (M' + M'')\,|\mathbf{y} - \mathbf{x}|,$$

para todo x e y en K, como se quería demostrar.

8.2 Teorema

Si S es un conjunto liso en $\mathcal{R}^n$, entonces S puede ser cubierto por un número finito de rectángulos coordenados de contenido total arbitrariamente pequeño. El recubrimiento puede ser hecho de modo tal que ningún punto de S esté en la frontera de la unión del conjunto de los rectángulos del recubrimiento.

Demostración. El caso en que S contiene un solo punto es trivialmente verdadero, de modo que podemos suponer $m \geq 1$. El conjunto liso S es la imagen bajo una función continuamente diferenciable $\mathcal{R}^m \xrightarrow{g} \mathcal{R}^n$, $m < n$, de un conjunto cerrado y acotado K en $\mathcal{R}^m$. Incluimos K en un cubo de lado s y subdividimos el cubo en cubos pequeños de lado s/N, donde N es un entero mayor que 1. Hay N^m de estos cubos. Por el Teorema 8.1, en cada uno de estos cubos que contiene puntos de K, tenemos que

$$\|g(\mathbf{x}) - g(\mathbf{y})\| \leq M\,\|\mathbf{x} - \mathbf{y}\| \leq M\,\frac{s}{N},$$

donde M es una constante que solo depende de K y g. Esto significa que la imagen por g de la parte de K en cada cubo está contenida en un cubo de lado $M(s/N)$. Entonces S está contenido en N^m cubos, cada uno de los cuales tiene un volumen $(Ms/N)^n$. El volumen total de estos cubos es a lo más $N^m(Ms/N)^n = (Ms)^m/N^{n-m}$. Como $n > m$, el volumen total puede hacerse tan pequeño como se quiera haciendo N suficientemente grande. Si hacemos igual a $(Ms + 1)/N$ el lado de cada rectángulo del recubrimiento, la última condición del teorema se satisface.

Como corolario tenemos que *todo conjunto liso tiene contenido cero.*

Ahora podemos demostrar el teorema de existencia para integrales que indicáramos al comienzo de la sección. Supongamos que f y B son como en la hipótesis. Debemos encontrar un número que sea la integral de Riemann de f en B. Para un reticulado arbitrario G que cubre a B, sea R_k el k-ésimo rectángulo acotado de G y sea f_k el *ínfimo* de f_B en R_k. Definimos

$$\underline{S}(G) = \sum_{k=1}^{N} \underline{f}_k V(R_k).$$

En forma análoga, definimos

$$\bar{S}(G) = \sum_{k=1}^{N} \bar{f}_k V(R_k),$$

donde $\bar{f}_k$ es el *supremo* de f_B en R_k. Es claro que

$$\underline{S}(G) \leq \sum_{k=1}^{N} (f_B(\mathbf{x}_k))V(R_k) \leq \bar{S}(G), \tag{2}$$

para una suma de Riemann arbitraria con respecto del reticulado G. Además, si G es el reticulado obtenido subdividiendo los rectángulos de G, entonces

$$\underline{S}(G) \leq \underline{S}(G') \leq \bar{S}(G') \leq \bar{S}(G).$$

En particular, si G y G'' son dos reticulados y G' contiene todos los rectángulos de ambas, entonces

$$\underline{S}(G) \leq \underline{S}(G') \leq \bar{S}(G') \leq \bar{S}(G''). \tag{3}$$

Definimos

$$I_B f = \textit{supremo de } \underline{S}(G),$$

donde el *supremo* se toma respecto de todos los reticulados G que cubren a B. De la relación (3), tenemos que

$$\underline{S}(G) \leq I_B f \leq \bar{S}(G)$$

de donde

$$-\bar{S}(G) \leq -I_B f \leq -S(G).$$

Si sumamos esta desigualdad con (2), obtenemos

$$\left| \sum_{k=1}^{N} (f_B(\mathbf{x}_k))V(R_k) - \bar{S}_B f \right| \leq \bar{S}(G) - \underline{S}(G),$$

en la cual la suma de Riemann es con respecto del reticulado G.

Todo lo que debemos hacer es mostrar que $\bar{S}(G) - S(G)$ puede hacerse arbitrariamente pequeño si tomamos G con malla suficientemente pequeña. Entonces, de acuerdo con la definición, habremos demostrado que la integral de f en B existe y es $I_B f$. Sea ϵ un número positivo. Por el Teorema 6.2, podemos cubrir la frontera de B, las superficies lisas que contienen los puntos de discontinuidad de f y cualquier otra superficie lisa en la cual no deseamos considerar los valores de f, con un número finito de rectángulos abiertos $R_1', \ldots, R_l'$ de medida total menor que ϵ. En la parte de B, no cubierta por estos rectángulos, f es continua; así, pues, por el Teorema 1.3, existe un $\delta > 0$ tal que $\bar{f}_k - \underline{f}_k < \epsilon$ sobre cualquier rectángulo R_k que pertenezca a una partición de norma menor que δ. Haciendo aún más pequeña la norma, por ejemplo, menor que δ', se puede llegar a una norma tal que todos los rectángulos $R_1', \ldots, R_l'$ estén contenidos siempre en un número finito de rectángulos $R_1'', \ldots, R_m''$ de toda partición con norma menor que δ' y tal que la medida total de estos últimos rectángulos sea menor que 2ϵ. Supóngase que los restantes rectángulos de una partición semejante G sean $R_1, \ldots, R_n$, que $|f| < M$ sobre B y que B esté contenido en un rectángulo de volumen C. Entonces

$$\begin{aligned}\bar{S}(G) - \underline{S}(G) &= \sum_{k=1}^{m} (\bar{f}_k'' - \underline{f}_k'')V(R_k'') + \sum_{k=1}^{n} (\bar{f}_k - \underline{f}_k)V(R_k) \\ &< (2M)(2\epsilon) + \epsilon C = \epsilon(4M + C).\end{aligned}$$

Así, pues, se ha hecho

$$\left| \sum_{k=1}^{N} f_B(\mathbf{x}_k) V(R_k) - I_B f \right| < \epsilon(4M + C)$$

para cualquier partición de norma suficientemente pequeña. Como ϵ se puede hacer arbitrariamente pequeño, queda terminada la demostración.

9. La fórmula del cambio de variables para integrales

En esta sección demostraremos el teorema del cambio de variable (Teorema 3.1) del Cap. 6.

Teorema

Sea $\mathcal{R}^n \xrightarrow{T} \mathcal{R}^n$ una transformación continuamente diferenciable. Sea R un conjunto en $\mathcal{R}^n$ cuya frontera está formada por un número finito de conjuntos lisos. Supongamos que R y su frontera están contenidos en el interior del dominio de T y que

1. T es inyectiva en R.
2. J, determinante jacobiano de T, es distinto de cero en todo R con la posible excepción de un número finito de conjuntos lisos.

Entonces, si la función f es acotada y continua en $T(R)$ (la imagen de R por T),

$$\int_{T(R)} f\, dV = \int_R (f \circ T)\, |J|\, dV.$$

Si f es discontinua en un conjunto liso contenido en R, entonces el teorema puede ser aplicado en el complemento de S en R. Al agregar S y $T(S)$ a los dominios de integración no se alteran las integrales, ya que estos conjuntos tienen contenido cero.

*Demostración**. Primero consideremos el caso especial en que f es la función constante 1 y T es lineal. Entonces, por el Teorema 3.9 del Cap. 2, salvo intercambios de variables, T puede ser expresada como un producto de transformaciones elementales de dos tipos: multiplicación numérica de una coordenada

$$M(x_1, \ldots, x_k, \ldots, x_n) = (x_1, \ldots, ax_k, \ldots, x_n); \tag{1}$$

y suma de un múltiplo de una coordenada a otra

$$A(x_1, \ldots, x_k, \ldots, x_n) = (x_1, \ldots, x_k + rx_j, \ldots, x_n). \tag{2}$$

* La demostración que ofrecemos aparece en J. Schwartz, «The formula for change of variable in a multiple integral», *American Math. Monthly*, vol. 61, núm. 2 (February, 1954). Véase también D. E. Varberg, «Changes of variables in multiple integrals», *American Math. Monthly*, vol. 78, número 1 (January, 1971).

Si miramos las matrices de estas transformaciones, es fácil ver que $\det M = a$ y $\det A = 1$. Si verificamos el teorema para estos casos especiales, aplicando sucesivamente la regla del producto para los determinantes, lo habremos demostrado para toda transformación lineal no singular. Sea R_k la proyección de R en el subespacio perpendicular al k-ésimo eje de coordenadas. Para cada punto $(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$ en R_k, sea I_k el conjunto de todos los x_k tales que $(x_1, \ldots, x_n)$ está en R. Para la transformación lineal (1), por integración iterada, tenemos que

$$\int_R |J|\, dV = \int_{R_k} dV_{n-1} \int_{I_k} |a|\, dx_k.$$

Si denotamos por $|a|\, I_k$ al conjunto de todos los números de la forma $|a|\, x_k$, donde x_k está en I_k, por un cambio de variables en una dimensión tenemos que

$$\int_{R_k} dV_{n-1} \int_{I_k} |a|\, dx_k = \int_{R_k} dV_{n-1} \int_{|a| I_k} du_k = \int_{M(R)} dV.$$

Para la transformación lineal (2), denotamos por $I_k + rx_j$ al conjunto de los números de la forma $x_k + rx_j$, donde x_k está en I_k. Entonces, por integración iterada y usando un cambio de variable en una dimensión, tenemos que

$$\int_R |J|\, dV = \int_R dV = \int_{R_k} dV_{n-1} \int_{I_k} dx_k = \int_{R_k} dV_{n-1} \int_{I_j + rx} dx_k = \int_{A(R)} dV.$$

Esto termina la demostración para el caso de transformaciones lineales T y funciones constantes f.

Para demostrar el teorema general, usaremos la siguiente norma para la matriz $A = (a_{ij})$ de una función lineal:

$$\|A\| = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|.$$

Si definimos análogamente una norma para vectores $\mathbf{x} = (x_1, \ldots, x_n)$ por

$$\|\mathbf{x}\| = \max_{1 \le j \le n} |\mathbf{x}_j|,$$

entonces es inmediato que

$$\|A\mathbf{x}\| \le \|A\|\, \|\mathbf{x}\|.$$

La razón para usar esta norma es que, si $\mathbf{x}$ e $\mathbf{y}$ son vectores en un cubo coordenado de lado de largo $2s$, entonces $\|\mathbf{x} - \mathbf{y}\| \le 2s$. Análogamente, si $\mathbf{p}$ es el centro de un cubo de ese tipo, entonces el cubo cerrado está formado por todos los vectores $\mathbf{x}$ tales que $\|\mathbf{x} - \mathbf{p}\| \le s$.

Ahora supongamos que C es un cubo de lado $2s$ y centro $\mathbf{p}$ contenido en R. Por el teorema del valor medio, para $\mathbf{x}$ en C,

$$T_k(\mathbf{x}) - T_k(\mathbf{p}) = T_k'(\mathbf{y}_k)(\mathbf{x} - \mathbf{p}),$$

donde T_k son las funciones coordenadas de T, e $\mathbf{y}_k$ es un punto en el segmento que une a $\mathbf{x}$ y $\mathbf{p}$. Entonces

$$\|T(\mathbf{x}) - T(\mathbf{p})\| \leq \max_{\mathbf{y} \text{ en } C} \|T'(\mathbf{y})\| \, \|\mathbf{x} - \mathbf{p}\| .$$

Esto implica que $T(C)$ está contenido en el cubo formado por los vectores $\mathbf{z}$ tales que

$$\|\mathbf{z} - T(\mathbf{p})\| \leq s \max_{\mathbf{y} \text{ en } C} \|T'(\mathbf{y})\| .$$

Tomando las h-ésimas potencias de ambo; miembros, tenemos que

$$V(T(C)) \leq \left\{ \max_{\mathbf{y} \text{ en } C} \|T'(\mathbf{y})\| \right\}^n V(C).$$

Notemos que si L es una transformación lineal inyectiva con matriz A y S tiene por frontera a un número finito de conjuntos lisos, entonces $V(L(S)) = |\det A| \, V(S)$. Esto se deduce del caso especial del teorema del cambio de variables para las transformaciones lineales que ya demostramos.

Ahora tomamos $S = T(C)$ y $A = [T'(\mathbf{x})]^{-1}$. Entonces, aplicando (3) con T remplazado por $(d_{\mathbf{x}}T)^{-1} \circ T$, tenemos que

$$\begin{aligned} |\det [T'(\mathbf{x})]^{-1}| \, V(T(C)) &= V\big((d_{\mathbf{x}}T)^{-1} \circ T(C)\big) \\ &\leq \left\{ \max_{\mathbf{y} \text{ en } C} \|[T'(\mathbf{x})]^{-1} T'(\mathbf{y})\| \right\}^n V(C) \end{aligned}$$

es decir,

$$V(T(C)) \leq |\det T(\mathbf{x})| \left\{ \max_{\mathbf{y} \text{ en } C} \|[T'(\mathbf{x})]^{-1} T'(\mathbf{y})\| \right\}^n V(C). \tag{4}$$

Dividamos el cubo C en un número finito de cubos $C_1, \ldots, C_N$ con centros $\mathbf{x}_1, \ldots, \mathbf{x}_N$ y tales que unos con otros cada par de ellos solo se intersecten en la frontera. Sea δ el máximo de las longitudes de los lados de los cubos. Aplicando (4) a cada C_k, tomando $\mathbf{x} = \mathbf{x}_k$ en cada caso y sumando, tenemos

$$V(T(C)) \leq \sum_{n=1}^{n} |\det T'(\mathbf{x}_k)| \left\{ \max_{\mathbf{y} \text{ en } C_k} \|[T'(\mathbf{x}_k)]^{-1} T'(\mathbf{y})\| \right\}^n V(C_k).$$

Como T es continuamente diferenciable, $T'(\mathbf{y})$ es una función continua de $\mathbf{y}$ y $[T'(\mathbf{x}_k)]^{-1} T'(\mathbf{y})$ se acerca a la matriz identidad cuando $\mathbf{y}$ tiende a $\mathbf{x}_k$. Entonces existe una función $h(\delta)$ que tiende a cero cuando δ tiende a cero, tal que

$$\left\{ \max_{\mathbf{y} \text{ en } C_k} \|[T'(\mathbf{x}_k)]^{-1} T'(\mathbf{y})\| \right\}^n \leq 1 + h(\delta).$$

Por tanto,

$$V(T(C)) \leq [1 + h(\delta)] \sum_{k=1}^{n} |\det T'(\mathbf{x}_k)| \, V(C_k).$$

Cuando δ se acerca a cero, la suma de la derecha se acerca a $\int_C |\det T'| \, dV$. Entonces la última desigualdad se transforma en

$$V(T(C)) \leq \int_C |\det T'| \, dV. \tag{5}$$

Una vez demostrada esta desigualdad, la usamos para demostrar la fórmula para conjuntos más generales que los cubos. Supondremos que $f \geq 0$. El caso general se obtiene considerando separadamente las partes positiva y negativa de f y sumando los resultados. Sea G un reticulado que cubre a R con malla δ. Sean $C_1, \ldots, C_N$ los cubos de G que están contenidos en R. Si R_N es la parte de R que no está contenida en ninguno de los cubos C_k, entonces $R = C_1 \cup \ldots \cup C_N \cup R_N$. Si $\mathbf{y}_k$ es un punto de C_k y $\mathbf{x}_k = T(\mathbf{y}_k)$, denotaremos $(f \circ T)\mathbf{y}_k = f(\mathbf{x}_k)$ por f_k. Entonces, usando (5), tenemos que

$$\sum_{k=1}^{N} f_k \int_{T(C_k)} dV \leq \sum_{k=1}^{N} f_k \int_{C_k} |J| \, dV.$$

De esto se tiene que

$$\begin{aligned} D &= \int_{T(R)} f \, dV - \int_R (f \circ T) \, |J| \, dV \\ &\leq \int_{T(R)} f \, dV - \sum_{k=1}^{N} f_k \int_{T(C_k)} dV + \sum_{k=1}^{N} f_k \int_{C_k} |J| \, dV - \int_R (f \circ T) \, |J| \, dV. \end{aligned}$$

Como $T(R) = T(C_1) \cup \ldots \cup T(C_N) \cup T(R_N)$, tenemos que

$$\begin{aligned} D \leq & \int_{T(R_N)} f \, dV + \sum_{k=1}^{N} \int_{T(C_k)} (f - f_k) \, dV \\ & + \sum_{k=1}^{N} \int_{C_k} (f_k - f \circ T) \, |J| \, dV - \int_{R_N} (f \circ T) \, |J| \, dV. \end{aligned}$$

Como T es continuamente diferenciable, del Teorema 8.1 de este Apéndice se deduce que existe una constante B para la cual

$$\|T(\mathbf{x}) - T(\mathbf{y})\| \leq B\|\mathbf{x} - \mathbf{y}\|, \qquad \text{para todo } \mathbf{x} \text{ e } y \text{ en } R \tag{6}$$

Ahora sea ϵ un número positivo. Como f es uniformemente continua en $T(R)$ (aplicar el Teorema 1.3 de este Apéndice a $T(R)$ y su frontera), podemos elegir la malla δ de G lo suficientemente pequeña como para que

$$|(f \circ T)\mathbf{y} - f_k| \leq \epsilon, \quad \text{para } \mathbf{y} \text{ en } C_k, \qquad k = 1, \ldots, N.$$

Usando (6) y, si es necesario, tomando δ aún más pequeño, tenemos que

$$|f(\mathbf{x}) - f_k| \leq \epsilon, \quad \text{para } \mathbf{x} \text{ en } T(C_k), \qquad k = 1, \ldots, N.$$

Entonces

$$D \leq \int_{T(R_N)} f\,dV - \int_{R_N} (f \circ T)\,|J|\,dV + \epsilon\{V(T(R)) + V(R)\}.$$

Como hemos supuesto que R tiene un volumen, existe un reticulado tal que $V(R_N) \leq \epsilon$. Usando (6) nuevamente y, si es necesario, disminuyendo nuevamente la malla, tenemos que $V(T(R_N)) \leq B^n\epsilon$. Entonces

$$D \leq \epsilon\{MB^n + M + V(T(R)) + V(R)\},$$

donde M es un número tal que $f \leq M$ en $T(R)$ y $(f \circ T)|J| \leq M$ en R. Como ϵ es arbitrario, $D \leq 0$, de modo que

$$\int_{T(R)} f\,dV \leq \int_R (f \circ T)\,|J|\,dV.$$

Si aplicamos esta última desigualdad al caso en que T se remplaza por T^{-1}, tenemos que

$$\int_R (f \circ T)\,|J|\,dV \leq \int_{T(R)} (f \circ T \circ T^{-1})\,|J \circ T^{-1}|\,|J^{-1}|\,dV.$$

donde J^{-1} es el determinante jacobiano de la transformación T^{-1}. Pero $(J \circ T^{-1})(J^{-1}) = 1$, de modo que

$$\int_{T(R)} f\,dV \leq \int_R (f \circ T)\,|J|\,dV \leq \int_{TR} f\,dV,$$

con lo cual se ha demostrado la igualdad deseada.

Si J es cero en un pedazo de superficie lisa S en R, entonces la demostración dada falla puesto que T^{-1} puede no ser continuamente diferenciable. Sin embargo, por el Teorema 8.2 de este Apéndice, S puede ser incluido en el interior de la unión U de un número finito de rectángulos de contenido arbitrariamente pequeño v, y la superficie imagen $T(S)$ está incluida, entonces, en la región imagen $T(U)$ cuyo contenido es a lo más $B^n v$, donde B es la constante de la relación (6). Entonces, aplicando la fórmula del cambio de variable a la región $R - U$, obtenemos

$$\left|\int_{T(R)} f\,dV - \int_R (f \circ T)\,|J|\,dV\right| \leq \left|\int_{T(U)} f\,dV\right| + \left|\int_U (f \circ T)\,|J|\,dV\right| \leq MB^n v + Mv,$$

donde $|f|$ y $|f \circ T|\,|J|$ son ambos menores que M. Cuando v tiende a cero, obtenemos la igualdad deseada.

Indice

Aceleración:
  magnitud de la, 199
  normal, 207
  tangencial, 207
Afín:
  función, 32, 240
  subespacio, 110
Amplitud, 151
Angulo, 151
Angulo de fase, 151
Angulo polar, 141
  forma polar, 141
Angulo sólido, 520
Aproximación:
  de Fourier, 383
  de Legendre, 419
  de Taylor, 366, 369, 371
  de tres puntos de Simpson, 480
  en promedio cuadrado, 416
Area:
  elemento de, 459
  superficie, 512

Base, 6, 12, 106
Base natural, 6
Cambio de variable, 449, 582
Camino aleatorio, 87
Campo:
  conservativo, 504
  de fuerzas, 212
  gradiente, 504
  vectorial o de vectores, 190
Campo incompresible, 536, 543
Campo irrotacional, 500, 536
Campo solenoidal, 543
Centro de masa, 210
Centroide, 448
Cerradura de un conjunto, 237
Circuito, 498
Circulación, 500, 525
Coeficientes:
  de Fourier, 153, 383, 414
  de Legendre, 418
  multinomiales, 371
Cofactor, 50, 66
Columna de una matriz, 15
Columna, fila asociada, 93
Columnas pivotales, 93
Combinación lineal, 6, 102
Complejo:
  complejo conjugado, 140

Complejo:
espacio vectorial, 137, 145
exponencial, 142
números, 139
producto interno, 148
suma compleja, 140
Componente centrípeta, 207
Componente de un vector, 38
Componente normal, 207
Componente tangencial, 207
Composición de funciones, 29, 115, 276
Conjunto abierto, 220, 236
Conjunto cerrado, 237
Conjunto de nivel, 187, 189
Conjunto liso, 435, 578
Conjunto ortonormal, 41, 152
Contenido, 433, 439
n-dimensional, 440
Convergencia:
convencional, 470
de series de Fourier, 385
de sucesiones, 253
puntual, 406
uniforme, 407
Convexo, 14
Coordenada, 38
Coordenada de un vector respecto de otro vector, 38
Coordenadas cilíndricas, 306
Coordenadas curvilíneas, 298, 303
Coordenadas de un vector, 170
Coordenadas esféricas, 301
bola en, 230
Coordenadas polares, 299
Coseno de dirección, 48
Criterio de Weierstrass, 408
Curva lisa, 326
por pedazos, 203, 385, 517
Curva rectificable, 203
Curvatura, 207

Definición explícita, 187, 189
de superficies, 329
Definición implícita, 187, 189, 290, 291
Definición paramétrica, 187, 189
Del, 536
Densidad de probabilidad, 449
Derivada, 195, 217, 242
direccional, 261, 262
Derivada exterior, 553
Derivada, diferenciación exterior, 553
Derivada parcial, 217
Derivada respecto a un vector, 244
Desarrollo:
de Fourier, 155, 390
de Legendre, 418
de Taylor, 366, 369, 376
multinomial, 371
Desarrollo de determinantes, 64
Desigualdad de Bessel, 415
Desigualdad de Cauchy-Schwarz, 39
Desigualdad triangular, 39, 149, 558
Determinante, 49, 50
desarrollo de un, 64
Determinante jacobiano, 284
Diferenciabilidad continua, 249
Diferencial, 241
de orden k, 371
Dimensión, 107
Dimensión finita, 108
Discontinuidad removible, 239
Divergencia de un campo, 500, 530
Dominio, espacio dominio, 1, 2, 189

Ecuación de calor, 396, 398
Ecuación de difusión, 398
Ecuación de ondas, 396
Ecuación de Laplace, 538
Ecuación de Poisson, 538
Ecuación homogénea, 95, 133
solución de la, 133
Ecuación, polinomio de Legendre, 418
Ecuación, raíces, vectores característicos, 160, 163, 357
Elemento de:
largo de arco, 206
superficie, 459
volumen, 459
Elemento de una matriz, 15
Elemento principal, 93
Energía cinética, 504
Energía potencial, 505
Equilibrio, 87
Escalar, 6
Espacio con producto interno, 39
Espacio imagen, 2
Espacio vectorial, 5
complejo, 139, 145
subespacio lineal, subespacio, 103

Exponencial, multiplicador exponencial complejo, 125

Familia convergente de subconjuntos, 466
Fila de una matriz, 15
Flujo, 500, 515, 548
Flujo estable, 226, 227
Flujo incompresible, 500
Forma cuadrática, 349
Forma de una matriz, 15
Forma diagonal, 350
Forma diferencial, 544, 545, 547
Forma escalonada, 102
Fórmula de Lagrange, 418
Fórmula del resto integral, 366
Fórmula de Stokes, 523
Fourier, aproximación de, 383, 419
coeficiente de, 153, 383, 414
desarrollo de, 155, 390
Frontera de una superficie, 517
Frontera de un conjunto, 237
Fuerza resultante, 86
Función:
afín, 32
bilineal simétrica, 349
compuesta, 29, 276
continua, 233
continuamente diferenciable, 249
diferenciable, 241
integrable, 434
inversa, 31, 118, 313
inyectiva, 28, 121
límite de, 194, 222, 230
ortogonal, 413
par, impar, 392
real, 182
traslación, 32
Función armónica, 535
Función, conjunto acotado, 433
Función de Cantor, 563
Función identidad, 287
Función propia, 416
Funciones lineales, 24, 114
composición de, 115
suma de, 114
Funciones coordenadas, 170

Gradiente, 266
campo, 504
de temperatura, 276, 519
Gradiente, local, 528
Gráfico, 182, 189

Homogenicidad conjugada, 148

Identidades de Green, 538
Identidades vectoriales, 537
Imagen, 2, 189
Ínfimo, 557
Integral condicionalmente convergente, 470
Integral de recta, 211
Integral, de superficie, 510, 513
de línea, 211
Integral de Riemann, 578
impropia, 466
Integral de una función, 434, 466
Integral iterada, 423, 424
Integral sumable en el sentido de Abel, 47
Integral vectorial, 201, 447
Integrales múltiples, 432
Isomorfismo, 173

Largo de arco, 563
elemento de, 206
parametrización por, 206
Largo de una curva, 203
Largo de un vector, 36
Lema de Riemann, 567
Ley del paralelogramo, 8
Límite de una función, 194, 220, 230
Linealmente dependiente, independiente, 11, 105

Magnitud de la aceleración, 199
Magnitud de una fuerza, 199
Malla, 433
Mapa coordenado, 173
Masa total, 205
Matriz, 15
de una transformación, 24
diagonal, 21
elemental, 99, 102
identidad, 19
inversa, 20
invertible, 20
jacobiana, 244
ortogonal, 41
producto de, 17
simétrica, 350
traspuesta, 41, 42

Matriz cero, 16
Matriz cuadrada, 15
Matriz de coeficientes, 75
Matriz reducida, 93
Máximo:
  absoluto, 388
  local, 388
  norma del, 558
Máximo, mínimo absoluto, 388
Menor, 49
Método de Newton, modificado, 252
Mínimo:
  absoluto, 388
  local, 388
Modificación elemental, 77
Momento angular, 201
Momento de inercia, 448
Momento lineal, 201
Multiplicación elemental, 77
Multiplicadores de Lagrange, 343
Multiplicación numérica, 140

Negativo definido, 364, 376
Norma, 39
  caja, del máximo, 559
Normal principal, 275
Normas equivalentes, 559
Núcleo, 103

Operaciones elementales, 77
  inversa, 78
  modificación, 77
  multiplicación, 77
  transformación, 99
  transposición, 78
Operador de Sturm-Liouville, 417
Operador lineal, 126, 416
  diferencial, 123
Operador simétrico, 539
Orientación:
  de la frontera de una superficie, 530
  de una superficie, 517
  negativa, 59
  positiva, 59, 517, 530, 531
Origen, 7

Paraboloide hiperbólico, 351
Paralela, 10, 110
Paralelepípedo, 58
Parametrización por el largo de arco, 206
Parametrizaciones equivalentes, 208, 519
Parte imaginaria, 140
Parte real, 140
Plano, 10, 110, 325
  por el origen, 110
Poligonalmente conexo, 265
Polinomio:
  cuadrático, 348
  de Legendre, 418
  de Taylor, 367, 369, 370
  homogéneo cuadrático, 349
  trigonométrico, 382
Polinomio característico, 130
Positivo definido, 351, 378
Potencial:
  logarítmico, 275
  newtoniano, 275, 543
Potencial de un campo, 504
Principio de superposición, 89, 134
Procedimiento de ortogonalización de Gram-Schmidt, 155
Producto cruz, 45, 60
Producto exterior, 548
Producto interno, 35
Producto interno complejo, 148
Producto punto, 35
Promedio, probabilístico, 472
Proyección, 7, 155
Punto aislado, 233
Punto crítico, 341
Punto de discontinuidad infinita, 464
Punto interior, 236
Punto límite, 230

Rango, 120
Rapidez, 198
Recta, 10, 110
  por el origen, 110
Rectángulo coordenado, 432
  abierto, 433
  cerrado, 432
Recubrimiento, 433
Red, 84
Región simple, 493, 531
Región simplemente conexa, 527
Regla de Cramer, 68
Regla de la cadena, 276, 278, 561
Regla de Leibnitz, 445
Regla del producto para determinantes, 57
Regla de Simpson, 481

Regla de Simpson, aproximación de Simpson por tres puntos, 480
Resto, de Taylor, 366
Reticulado, 433
Rotor, 499, 521

Segmento de recta, 14
Separación de variables, 398
Serie de Fourier, 382, 385
  convergencia, 565
Simetría, 164
  con respecto de un producto interno, 416
Simetría conjugada, 148
Simetría hermítica, 168
Similar, 176
Sistemas de ecuaciones lineales, 75
Sistemas equivalentes de ecuaciones lineales, 76
Solución particular, 133
Subespacio, 103
  afín, 110
Subespacio generado, 104
Subespacio lineal, 103
Sucesión de Cauchy, 558
Sucesión ortonormal, 413
Suma:
  de funciones, 114
  de matrices, 16
  de n-tuplas, 4
Suma de Riemann, 434
Superficie:
  área, 512
  cuadrática, 361
  definida explícitamente, 329
  definida implícitamente, 331
  de revolución, área, 520
  equipotencial, 275
  integral de, 510, 513
  lisa, 326, 511
Superficie, 325
  pedazo de, 511
Superficie cerrada, 526
Superficie lisa por pedazos, 517
Superficie orientable, 517
Supremo, 557

Tangente:
  plano, 270, 325
  recta, 197, 242, 270
  vector, 197
Teorema de Gauss, 500, 530, 531
Teorema de Green, 490, 493
Teorema de la divergencia, 500
Teorema de la función implícita, 319, 571
Teorema de la función inversa, 315, 570
Teorema de Lagrange, 573
Teorema de linealidad, 441
Teorema del valor medio, 264
Teorema de positividad, 441
Teorema de Stokes, 499, 521, 523
Teorema de Taylor, 373, 576
Teorema fundamental del cálculo, 272
Torque, 201
Trabajo, 213
Transformación:
  aguda, 167
  elemental, 99
Transformación de cizalla, 167
Transformación lineal simétrica, 366
Transformación lineal y afín, 33
Transposición elemental, 78
Traslación, 238
  de funciones, 32
Trayectoria, 226

Valor absoluto, 140
Valor máximo local mínimo local, 338
Valor principal, 200
Valor, punto extremo, 338, 376
Valores propios, 161, 357, 416
Varianza, 472
Vecindad, 236
Vector, 5
  largo de un, 36
  velocidad de un, 198
Vector aceleración, 199
Vector binormal, 200
Vector columna, de dimensión n, 15
Vector coordenado, 170
Vector fila de dimensión n, 15
Vector fuerza, 199, 201
  campo de, 212
Vector normal, 270
Vector unitario, 261
Vectores dependientes, 11, 105
Vectores independientes, 11, 105
Vectores ortogonales, 41
Vectores propios, 160, 357, 416
Volumen, 433, 439
  orientado, 62